

STIC-Biotech/ChemLib

99925

From: Chan, Christina
Sent: Monday, July 28, 2003 6:04 PM
To: Sullivan, Daniel; STIC-Biotech/ChemLib
Subject: RE: RUSH Sequence Search for 09834291

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Sullivan, Daniel
Sent: Monday, July 28, 2003 1:45 PM
To: Chan, Christina
Subject: RUSH Sequence Search for 09834291
Importance: High

Hi Chris-

Could you please approve the following search for an after final amended case.

Please search for the following in the commercial, interference and issued patent databases:

A nucleic acid sequence comprising SEQ ID NO: 1, 2, 3, 4, 6, 15 or 24;

A nucleic acid sequence consisting of SEQ ID NO: 10, 12 or 14

Thank you.

Daniel M. Sullivan
Examiner AU 1636
Room: 12D12
Mail Box: 11E12
Tel: 703-305-4448

Searcher: Jan
Phone: 4657
Location: _____
Date Picked Up: 7/29/03
Date Completed: 8/3/03
Searcher Prep/Review: _____
Clerical: 20
Online time: 15

TYPE OF SEARCH: ☒
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: *
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Sun Aug 3 09:03:33 2003

us-09-834-291-1.rge

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 12552.1 Seconds
(without alignment)
10468.541 Million cell updates/sec

Title: US-09-834-291-1

Sequence: 1 tgcagactctcagatgataatg.....ttgcagatgagtaacaaag 3212

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481336 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1.*

- 1: gb_da.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_ov.*
- 7: gb_ov.*
- 8: gb_ov.*
- 9: gb_ov.*
- 10: gb_ov.*
- 11: gb_ov.*
- 12: gb_ov.*
- 13: gb_ov.*
- 14: gb_ov.*
- 15: gb_ov.*
- 16: gb_ov.*
- 17: gb_ov.*
- 18: gb_ov.*
- 19: gb_ov.*
- 20: gb_ov.*
- 21: gb_ov.*
- 22: gb_ov.*
- 23: gb_ov.*
- 24: gb_ov.*
- 25: gb_ov.*
- 26: gb_ov.*
- 27: gb_ov.*
- 28: gb_ov.*
- 29: gb_ov.*
- 30: gb_ov.*
- 31: gb_ov.*
- 32: gb_ov.*
- 33: gb_ov.*
- 34: gb_ov.*
- 35: gb_ov.*
- 36: gb_ov.*
- 37: gb_ov.*
- 38: gb_ov.*
- 39: gb_ov.*
- 40: gb_ov.*
- 41: gb_ov.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match length	DB ID	Description
1	3212	100.0	3212 6	AX026089
2	3193	99.4	45121 6	AX026089
3	3185	99.2	187313 9	AX026089
4	2346.6	73.1	2827 6	AX026089
5	2211.8	68.7	2344 9	HSCD955FR
6	2143.4	68.7	2165 9	HUMFAS
7	1899.6	59.1	1877 9	HSAP1
8	1850.8	57.6	1608 9	HSFASX1
9	1536.8	47.8	1608 6	AX026091
10	957.4	29.9	1608 6	AX026091
11	957.4	22.2	720 6	HSN278012
12	713	22.2	702 9	HSN278011
13	698.8	21.7	702 9	HSN278013
14	697.2	21.7	702 9	HSN278013
15	695.6	13.4	436 9	HSN278013
16	431.2	13.4	436 9	HSN278013
17	429.6	10.6	266 6	HSN278013
18	342	8.2	266 6	HSN278013
19	262.8	8.2	266 6	HSN278013
20	257	8.0	398 11	G27038
21	257	8.0	251 6	AR143111
22	252	7.8	251 6	AX587673
23	252	7.8	251 6	AX587673
24	252	7.8	251 6	AX587673
25	252	7.8	251 6	AX587673
26	252	7.8	251 6	AX587673
27	226	7.0	2471 6	HSAP1
28	226	7.0	2471 6	HSAP1
29	226	7.0	2534 6	AR163572
30	226	7.0	2534 6	AR163572
31	226	7.0	2534 6	AR163572
32	226	7.0	2534 6	AR163572
33	226	7.0	2534 6	AR163572
34	226	7.0	2534 6	AR163572
35	208.6	6.5	1728 4	HSAP1
36	78.6	2.5	1728 4	HSAP1
37	78.6	2.5	1728 4	HSAP1
38	78.6	2.5	1728 4	HSAP1
39	78.6	2.5	1728 4	HSAP1
40	78.6	2.5	1728 4	HSAP1
41	78.6	2.5	1728 4	HSAP1

ALIGNMENTS

RESULT 1
LOCUS AX026089
DEFINITION Sequence 1 from Patent DE19847779.
ACCESSION AX026089
VERSION AX026089.1 GI:10187520
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

Pred. No. is the number of results predicted by chance to have a

	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
BASE COUNT	778 a	841 t
ORIGIN	784 c	809 g

	0;	Mismatches	0;	Indels	0;	Gaps	0
1	1	0	0	0	0	0	0

1 GAGGAGCTCTGAGGAAATATGCTGGTAAATAATAAAAATAACTTTAGAGATGCCAAACTGT 60
61 TTTCCCGAAGAACACACGACATTCATTATGGGCTTCATTCAATAGATTCCTCAAGATTC 60
61 TTTCCCGAAGAACACGACATTCATTATGGGCTTCATTCAATAGATTCCTCAAGATTC 60
121 AAGCAAAAGAAAGTTTGGGGAACAGATATATATAATTACCAACCCCTTGACATTGCA 120
121 AAGCAAAAGAAAGTTTGGGGAACAGATATATATAATTACCAACCCCTTGACATTGCA 120
181 TAAGGAGCCCTGAGAGTTTGGATTAAAGAAAGTTTCAATTAAAGTACCCAGATTTT 180
181 TAAGGAGCCCTGAGAGTTTGGATTAAAGAAAGTTTCAATTAAAGTACCCAGATTTT 180
241 CTAAAGATTATTTGACCATAGAAACATATGTCTCCCGCAACAGACATATTTCTATGCTT 240
241 CTAAAGATTATTTGACCATAGAAACATATGTCTCCCGCAACAGACATATTTCTATGCTT 240
301 GAACCTGAGATATATAGACGTATAGGGGTAGAGGTAGGGGAAAGGGGGTATGCGATAGA 300
301 GAACCTGAGATATATAGACGTATAGGGGTAGAGGTAGGGGAAAGGGGGTATGCGATAGA 300
361 AAGACAGAGACCTTGGGAGCAAGAAATATCTAAGTTAATTCGATCTGCTATTTATTA 360
361 AAGACAGAGACCTTGGGAGCAAGAAATATCTAAGTTAATTCGATCTGCTATTTATTA 360
421 ACTAACCATCTTTGCGCAATGTGCTTAAGCTTTTGGGCTACATTTTATTTATTTAAG 420
421 ACTAACCATCTTTGCGCAATGTGCTTAAGCTTTTGGGCTACATTTTATTTATTTAAG 420
481 TAAGTTAAATATACATCTCATCTGAGGCTATATATGAATTAAGTAAAGAAATGTC 480
481 TAAGTTAAATATACATCTCATCTGAGGCTATATATGAATTAAGTAAAGAAATGTC 480
541 CACATATGTAGATTGCTGGCTTAATATATACATCAAGAGAAATGATATTTGCTCAATGT 540
541 CACATATGTAGATTGCTGGCTTAATATATACATCAAGAGAAATGATATTTGCTCAATGT 540
601 CTTTCCCTCTTTTCTCTCTTCCCTCTTCCATCTCTTCCCTTCACTCTCTCTTTC 600
601 CTTTCCCTCTTTTCTCTCTTCCCTCTTCCATCTCTTCCCTTCACTCTCTCTTTC 600
661 CTTCCCGCAACCCCTTTTCTCTCTCTTTTCAATTTTATTTATTTAAATGAATTTTC 660
661 CTTCCCGCAACCCCTTTTCTCTCTCTTTTCAATTTTATTTATTTAAATGAATTTTC 660
721 AATTGGAATATTTAGGATTTGAAAAAATTTGACAGAGATAAAGAGAAATGGCCATA 720
721 AATTGGAATATTTAGGATTTGAAAAAATTTGACAGAGATAAAGAGAAATGGCCATA 720
781 TACCATCTCTCTTATCCCACTCTCTTTTGTGTCTAATAGATGTCTCAGAGTGTGTCACA 780
781 TACCATCTCTCTTATCCCACTCTCTTTTGTGTCTAATAGATGTCTCAGAGTGTGTCACA 780
841 GCGTGGACAGCCGAGGCTTCTCTCATGGCACTAACAGTCTATCTAAAGGTGAGAACAGAG 840
841 GCGTGGACAGCCGAGGCTTCTCTCATGGCACTAACAGTCTATCTAAAGGTGAGAACAGAG 840
900 GAGCTGGACAGCCGAGGCTTCTCTCATGGCACTAACAGTCTATCTAAAGGTGAGAACAGAG 900

QY 901 ACAAGCTTATCAACACCTTCAAGACTGTGTGTAAGTGTGACGTGACAGATGTGCAAAACACAGG 960
Db 901 ACAAGCTTATCAACACCTTCAAGACTGTGTGTAAGTGTGACGTGACAGATGTGCAAAACACAGG 960
QY 961 GTAATGGAAGAGCCCTCAGAGAGGTAACTTAACTCAAGATTGAGAGGCCCAACAGGCTCCAG 960
Db 961 GTAATGGAAGAGCCCTCAGAGAGGTAACTTAACTCAAGATTGAGAGGCCCAACAGGCTCCAG 960
QY 1021 AAGAAATGTCACTGAGAGAGAGCCGGAAGATGGAACAGTGGAGCTTAAGCAAAAGGTTAT 1080
Db 1021 AAGAAATGTCACTGAGAGAGAGCCGGAAGATGGAACAGTGGAGCTTAAGCAAAAGGTTAT 1080
QY 1081 TAAATGTATTAAATGAGGTTAAATCTAAATGAGAGAGGAGACTTGCAGAGTGAAGGTG 1140
Db 1081 TAAATGTATTAAATGAGGTTAAATCTAAATGAGAGAGGAGACTTGCAGAGTGAAGGTG 1140
QY 1141 CAGAGCTGTGTGAGAGATGCAAGAAAGAAATCTGAAACCTTAAATGTGTCACTGTGAAC 1200
Db 1141 CAGAGCTGTGTGAGAGATGCAAGAAAGAAATCTGAAACCTTAAATGTGTCACTGTGAAC 1200
QY 1201 TGCATTCGAATTCAGATTCAAGATCTACGTTCCAAAGCAATGACCTTGTGTCACATGTGAAC 1260
Db 1201 TGCATTCGAATTCAGATTCAAGATCTACGTTCCAAAGCAATGACCTTGTGTCACATGTGAAC 1260
QY 1261 ATGCTAACTACCTAAGAGAGTATCTACGTTCCAAAGCAATGAGATCTTGTGAATTC 1320
Db 1261 ATGCTAACTACCTAAGAGAGTATCTACGTTCCAAAGCAATGAGATCTTGTGAATTC 1320
QY 1321 CACCAGAGCAGCAAGAAATTAACAAGATTTTTTTTAAAGAAATTTGGCAGCAATATATG 1380
Db 1321 CACCAGAGCAGCAAGAAATTAACAAGATTTTTTTTAAAGAAATTTGGCAGCAATATATG 1380
QY 1381 AGTAAAGAGAGCAGCAAGATTAATGTGAAGTTTAAATATAGCTGAGGAGCTATGCAATTTGG 1440
Db 1381 AGTAAAGAGAGCAGCAAGATTAATGTGAAGTTTAAATATAGCTGAGGAGCTATGCAATTTGG 1440
QY 1441 CTAAAGTTGTAGCTTTGTTTCTCTTGAAGAAATTAATACTAAGAGGAGCCCTCCCTTTTC 1500
Db 1441 CTAAAGTTGTAGCTTTGTTTCTCTTGAAGAAATTAATACTAAGAGGAGCCCTCCCTTTTC 1500
QY 1501 AGAGCCTTAAGGCGCAACATCTGTACTTTTCAATATGAGTTAATCTGTCAATTCAGAAAG 1560
Db 1501 AGAGCCTTAAGGCGCAACATCTGTACTTTTCAATATGAGTTAATCTGTCAATTCAGAAAG 1560
QY 1561 TCTGTAGAGCTCTCATGTTGAGCCACAACATGGAACAGGCCAGTCAAAATGCCCGCAAGT 1620
Db 1561 TCTGTAGAGCTCTCATGTTGAGCCACAACATGGAACAGGCCAGTCAAAATGCCCGCAAGT 1620
QY 1621 CTATTCTGAGTGAATCTCAGCAATTAAGCCAGGCTCTGTACCCAGGAGAGACTCTGCG 1680
Db 1621 CTATTCTGAGTGAATCTCAGCAATTAAGCCAGGCTCTGTACCCAGGAGAGACTCTGCG 1680
QY 1681 CTGTGAGCTCATTTCTCTTGAAGACTTCTCCCAACTTCCAGGTTGAATCAAGCAGAAG 1740
Db 1681 CTGTGAGCTCATTTCTCTTGAAGACTTCTCCCAACTTCCAGGTTGAATCAAGCAGAAG 1740
QY 1741 CTTTGAAGAAAGGCGAGAGGCGGCTCTGAGAGCTTCACTGAATGAGATGCAAGC 1800
Db 1741 CTTTGAAGAAAGGCGAGAGGCGGCTCTGAGAGCTTCACTGAATGAGATGCAAGC 1800
QY 1801 ACTGAGAGAAAGCCCGGAGCAGGAATGCCATTTGTGCAAGAACTCTGACTCTTCT 1860
Db 1801 ACTGAGAGAAAGCCCGGAGCAGGAATGCCATTTGTGCAAGAACTCTGACTCTTCT 1860
QY 1861 CACCTGAATCTTCTCCCTCTCCCTAACCAGGCGCAGGCAAGTGTGAATCAATGAGGCC 1920
Db 1861 CACCTGAATCTTCTCCCTCTCCCTAACCAGGCGCAGGCAAGTGTGAATCAATGAGGCC 1920
QY 1921 CTCCCAACCCGGGAGTTTCCCAAGAGGCTTCTTCTCCATCTCTTGAACAACGGAGCT 1980
Db 1921 CTCCCAACCCGGGAGTTTCCCAAGAGGCTTCTTCTCCATCTCTTGAACAACGGAGCT 1980

us-09-834-291-1.rge

[illegible]

[illegible]

[illegible]

Accession	Sequence	Length
QY	902	
QY	1018	1019
Db	1018	1019
QY	961	1137
QY	1078	1079
Db	1078	1079
QY	1020	1197
QY	1138	1139
Db	1138	1139
QY	1257	1257

[illegible]

QY 1620 TCTTCTCTGAGTACTCCAGCAATTAGCCAAAGGCTCTGTACCAGGCAAGACTCTGCG 1679
 Db 1738 TCCTTCTCTGAGTACTCCAGCAATTAGCCAAAGGCTCTGTACCAGGCAAGACTCTGCG 1679
 QY 1680 GCTCTGAGCTCCCAATTCTCTCTTCAAGACTCCCCCACTTCCAGGTGTAACTACAGCAGAA 1739
 Db 1798 GCTCTAGCTCCCAATTCTCTCTTCAAGACTCCCCCACTTCCAGGTGTAACTACAGCAGAA 1739
 QY 1740 GCTTTAGAAAGGGGAGAGAGCGGCTCTCAGAGTCTCACTTAAATGTAGCAATGCCAACG 1857
 Db 1858 GCTTTAGAAAGGGGAGAGAGCGGCTCTCAGAGTCTCACTTAAATGTAGCAATGCCAACG 1799
 QY 1800 CACTGCAGAAAGCGCCCGGAGCAGAAATGCCCAATTGTGACAGAACCTTGACTCTCTCC 1859
 Db 1918 CACTGCAGAAAGCGCCCGGAGCAGAAATGCCCAATTGTGACAGAACCTTGACTCTCTCC 1859
 QY 1860 TCACCCGACTTCTCCCTCTCCCTACCCGCGGAGGCCAAAGTTGCTGAATCAATGAGAC 1919
 Db 1978 TCACCCGACTTCTCCCTCTCCCTACCCGCGGAGGCCAAAGTTGCTGAATCAATGAGAC 1919
 QY 1920 CCTCCCCCAACCCGGGAGTCCCCACAGAGCTTCTTCCACTCTCTCCAGCAACCGGGG 2037
 Db 2038 CCTCCCCCAACCCGGGAGTCCCCACAGAGCTTCTTCCACTCTCTCCAGCAACCGGGG 1979
 QY 1980 TTTTGTGAGCTCGCTCTGATCTCGCGCAAGTGCACACAGGTGTCAAAAGACGCTT 2157
 Db 2098 TTTTGTGAGCTCTCTCTGATCTCGCGCAAGTGCACACAGGTGTCAAAAGACGCTT 2039
 QY 2040 CTGGGAGGTAGGGAAGCGGTTTACAGTAGTACTGTGGCTGAGACCTCAGAGGGGCGGCACTG 2157
 Db 2158 CTGGGAGGTAGGGAAGCGGTTTACAGTAGTACTGTGGCTGAGACCTCAGAGGGGCGGCACTG 2099
 QY 2100 GCAACGGAAACAACCTCGAGGCGAGCCCTGGCTGCCAGCGGAGACTGCTTCTCCCGC 2159
 Db 2218 GCAACGGAAACAACCTCGAGGCGAGCCCTGGCTGCCAGCGGAGACTGCTTCTCCCGC 2159
 QY 2160 GGGTTGTGAGCCCGCTCAATACGAATGTGGGAGAACTCTTCACTTCGAGAGATTGCTC 2219
 Db 2278 GGGTTGTGAGCCCGCTCAATACGAATGTGGGAGAACTCTTCACTTCGAGAGATTGCTC 2219
 QY 2220 AACCAAC 2226
 Db 2338 AACCAAC 2344

	protein_bind	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/clone="pF7"	
		/tissue_type="blood"	
	79..87		
	/bound_molecly="NF-IL6"		
	276..284		
	/bound_molecly="NF-IL6"		
	481..489		
	/bound_molecly="NF-IL6"		
	772..780		
	/bound_molecly="NF-IL6"		
	934..940		
	/bound_molecly="NF-IL6"		
	948..955		
	/bound_molecly="AP-1"		
	1298..1306		
	/bound_molecly="Ets"		
	1436..1483		
	/bound_molecly="NF-IL6"		
	1436..1813		
	/number=1		
	1436..1781		
	1782..1813		
	/codon_start=1		
	/product="Pae antigen"		
	/protein_id="BA20850.1"		
	/db_xref="GI:443150"		
	/translation="MGITWTLPLV"		
	1814..2165		
	/number=1		
BASE COUNT	508 a	555 c	555 g
ORIGIN			547 t

Query Match	66.7%;	Score 2143.4;	DB 9;	Length 2165;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 2158;	Conservative	0;	Mismatches	6;
			Indels	1;
			Gaps	1;
Qy	447	AAGCTTTTGGCTACATTTTATTTTATTTGTAAGTAAGTTAATATATATCACTATCTACT	506	
Db	1	AAGCTTTTGGCTACATTTTATTTTATTTGTAAGTAAGTTAATATATATCACTATCTACT	506	
Qy	507	GGGCTTAATAGTAAAGTAAAGAAAGATCCACATATGAGTGGCTGGCTTAA	566	
Db	61	GGGCTTAATAGTAAAGTAAAGAAAGATCCACATATGAGTGGCTGGCTTAA	566	
Qy	567	TTCACTCAAGAGATCTGATTTTGTCAATGTCTTCCCTTTTTCCTCTCC	626	
Db	121	TTCACTCAAGAGATCTGATTTTGTCAATGTCTTCCCTTTTTCCTCTCC	626	
Qy	627	TCTTTCCATTCCTCTTCCCTTACCTGCTGCTTCCCTTCCCTTCCCTCTCT	180	
Db	181	TCTTTCCATTCCTCTTCCCTTACCTGCTGCTTCCCTTCCCTTCCCTCTCT	180	
Qy	687	TCCTTTTAACTTTTATTTTAAATGAACCTTTTCAATTTGGAATGTTTGAATTTCA	746	
Db	241	TCCTTTTAACTTTTATTTTAAATGAACCTTTTCAATTTGGAATGTTTGAATTTCA	746	
Qy	747	AAATTTGACAGATTAATTAACAGAGATGCCATTAACAATCTCTTAATCCACTTTT	806	
Db	301	AAATTTGACAGATTAATTAACAGAGATGCCATTAACAATCTCTTAATCCACTTTT	806	
Qy	807	TTGTGTCAATTAAGTGTCAAGTGTGTGCAACAAAGCTGGACGCCAGAGGTTTCTTA	866	
Db	361	TTGTGTCAATTAAGTGTGTGCAACAAAGCTGGACGCCAGAGGTTTCTTA	866	
Qy	867	TTGGACTTAACAGTCTCTGAAAGGTGGAAACAGAGACAAGGCTTAAACCTTAACAAGT	926	
Db	421	TTGGACTTAACAGTCTCTGAAAGGTGGAAACAGAGACAAGGCTTAAACCTTAACAAGT	926	
Qy	927	GGTGTAACTGACAGATGCAAAACACAGAGGATGGAAGACCTCTCAGAGGGTTAA	966	
Db	481	GGTGTAACTGACAGATGCAAAACACAGAGGATGGAAGACCTCTCAGAGGGTTAA	966	
			540	

987 CCTAACCTTGAATTTGAGGGGCC-AAACAGGCTCCAGAGAAATATCTCACTGAGAGGAAGC 1045
Db CCTAACCTTGAATTTGAGGGGCCAAACAGGCTCCAGAGAAATATCTCACTGAGAGGAAGC 600
Qy CTGAAGATGACAGTGGGCTAAGCAAAAGGTTATTAATGTTATTAATGGATTGAATC 1105
Qy CTGAAGATGACAGTGGGCTAAGCAAAAGGTTATTAATGTTATTAATGGATTGAATC 660
Db CTGAAGATGACAGTGGGCTAAGCAAAAGGTTATTAATGTTATTAATGGATTGAATC 1165
Qy TAAATGGGAAGGAGAGGTTGCAGAGTGAAGGTGCAGAGCTTGGTGGACATGCGCAAG 1106
Db TAAATGGGAAGGAGAGGTTGCAGAGTGAAGGTGCAGAGCTTGGTGGACATGCGCAAG 720
Qy TAAATGGGAAGGAGAGGTTGCAGAGTGAAGGTGCAGAGCTTGGTGGACATGCGCAAG 661
Qy GAATACCTAAACCTTGAATGTTGCTCAAGTCTGGAATCTCATCAATATTCAGGTTCAATAT 1166
Db GAATACCTAAACCTTGAATGTTGCTCAAGTCTGGAATCTCATCAATATTCAGGTTCAATAT 720
Qy GATGCTATATCAACATACCTCTCTGTAATTAATCAAGTCAATCTCAAGATCAATCT 1225
Db GATGCTATATCAACATACCTCTCTGTAATTAATCAAGTCAATCTCAAGATCAATCT 840
Qy ACCGTTCCAAAGCAATGATGATTTGAACAGTGTCAACAGAGCAAGAAATTAACAG 1286
Db ACCGTTCCAAAGCAATGATGATTTGAACAGTGTCAACAGAGCAAGAAATTAACAG 900
Qy ACCGTTCCAAAGCAATGATGATTTGAACAGTGTCAACAGAGCAAGAAATTAACAG 841
Qy ATTTTTTTTTTAAAGAAATTTGGCAGAGAAATATGATATGACAGAGCAAGAGATATG 1346
Db ATTTTTTTTTTAAAGAAATTTGGCAGAGAAATATGATATGACAGAGCAAGAGATATG 960
Qy ATTTTTTTTTTAAAGAAATTTGGCAGAGAAATATGATATGACAGAGCAAGAGATATG 901
Db ATTTTTTTTTTAAAGAAATTTGGCAGAGAAATATGATATGACAGAGCAAGAGATATG 1405
Qy TGAATGTTATATATGCTGGGGCTATGCGATTTGGCTTAATGTTAGTTGTTGTTCT 1406
Db TGAATGTTATATATGCTGGGGCTATGCGATTTGGCTTAATGTTAGTTGTTGTTCT 1020
Qy TGAATGTTATATATGCTGGGGCTATGCGATTTGGCTTAATGTTAGTTGTTGTTCT 961
Db TGAATGTTATATATGCTGGGGCTATGCGATTTGGCTTAATGTTAGTTGTTGTTCT 1525
Qy CTGGAGAAATTAATTAATGAGGGGCTCCCTTTTCAGAGCTTAATGGGCGCAATCTGTA 1466
Db CTGGAGAAATTAATTAATGAGGGGCTCCCTTTTCAGAGCTTAATGGGCGCAATCTGTA 1080
Qy CTGGAGAAATTAATTAATGAGGGGCTCCCTTTTCAGAGCTTAATGGGCGCAATCTGTA 1021
Db CTGGAGAAATTAATTAATGAGGGGCTCCCTTTTCAGAGCTTAATGGGCGCAATCTGTA 1585
Qy CTTTTTCATATGATTTAACTGTCCTCAATCCAGAAAGCTCTGAGAGCTCTCATGTTGAGGC 1526
Db CTTTTTCATATGATTTAACTGTCCTCAATCCAGAAAGCTCTGAGAGCTCTCATGTTGAGGC 1140
Qy CTTTTTCATATGATTTAACTGTCCTCAATCCAGAAAGCTCTGAGAGCTCTCATGTTGAGGC 1081
Db CTTTTTCATATGATTTAACTGTCCTCAATCCAGAAAGCTCTGAGAGCTCTCATGTTGAGGC 1645
Qy ACAACATGACAGCCAGTCAATATGCCCCGCAAGCTCTTCTCTGAGAGCTCCAGCAATTT 1586
Db ACAACATGACAGCCAGTCAATATGCCCCGCAAGCTCTTCTCTGAGAGCTCCAGCAATTT 1200
Qy ACAACATGACAGCCAGTCAATATGCCCCGCAAGCTCTTCTCTGAGAGCTCCAGCAATTT 1141
Db ACAACATGACAGCCAGTCAATATGCCCCGCAAGCTCTTCTCTGAGAGCTCCAGCAATTT 1705
Qy AGCCAGAGCTCTGTAACCCAGAGAGAGCTCTGAGAGCTCTGAGCTCTCATGTTGAGGC 1646
Db AGCCAGAGCTCTGTAACCCAGAGAGAGCTCTGAGAGCTCTGAGCTCTCATGTTGAGGC 1260
Qy AGCCAGAGCTCTGTAACCCAGAGAGAGCTCTGAGAGCTCTGAGCTCTCATGTTGAGGC 1201
Db AGCCAGAGCTCTGTAACCCAGAGAGAGCTCTGAGAGCTCTGAGCTCTCATGTTGAGGC 1765
Qy CTTCCCCCAATCTCCAGGTTGAACCTACAGAGAAAGCTTTAAGAAAGGAGAGAGCGGCG 1706
Db CTTCCCCCAATCTCCAGGTTGAACCTACAGAGAAAGCTTTAAGAAAGGAGAGAGCGGCG 1320
Qy CTTCCCCCAATCTCCAGGTTGAACCTACAGAGAAAGCTTTAAGAAAGGAGAGAGCGGCG 1261
Db CTTCCCCCAATCTCCAGGTTGAACCTACAGAGAAAGCTTTAAGAAAGGAGAGAGCGGCG 1825
Qy TTTGAGAGTCTCACTGTAAGTGAATGCAATGCAAGCTCACTGAGAGAGAGCGCCCGGACAGGA 1766
Db TTTGAGAGTCTCACTGTAAGTGAATGCAATGCAAGCTCACTGAGAGAGAGCGCCCGGACAGGA 1380
Qy TTTGAGAGTCTCACTGTAAGTGAATGCAATGCAAGCTCACTGAGAGAGAGCGCCCGGACAGGA 1321
Db TTTGAGAGTCTCACTGTAAGTGAATGCAATGCAAGCTCACTGAGAGAGAGCGCCCGGACAGGA 1885
Qy ATGCCATTTGTCGAAGAACCCGAGTCTCTCTCAACCTGACCTTCTCCCTCCCTAC 1826
Db ATGCCATTTGTCGAAGAACCCGAGTCTCTCTCAACCTGACCTTCTCCCTCCCTAC 1440
Qy ATGCCATTTGTCGAAGAACCCGAGTCTCTCTCAACCTGACCTTCTCCCTCCCTAC 1381
Db ATGCCATTTGTCGAAGAACCCGAGTCTCTCTCAACCTGACCTTCTCCCTCCCTAC 1945
Qy CCGCGGAGAGCCAAATTTGCTGAATCAATGAGAGAGCTCCCAACCCGAGGCTTCCAGC 1886
Db CCGCGGAGAGCCAAATTTGCTGAATCAATGAGAGAGCTCCCAACCCGAGGCTTCCAGC 1500
Qy CCGCGGAGAGCCAAATTTGCTGAATCAATGAGAGAGCTCCCAACCCGAGGCTTCCAGC 1441
Db CCGCGGAGAGCCAAATTTGCTGAATCAATGAGAGAGCTCCCAACCCGAGGCTTCCAGC 2005
Qy GAGGCTCTCTCCATCTCTGACCAACCGGAGCTTTTCTGAGAGTGTCTCTGATCTCG 1946
Db GAGGCTCTCTCCATCTCTGACCAACCGGAGCTTTTCTGAGAGTGTCTCTGATCTCG 1560
Qy GAGGCTCTCTCCATCTCTGACCAACCGGAGCTTTTCTGAGAGTGTCTCTGATCTCG 1501
Db GAGGCTCTCTCCATCTCTGACCAACCGGAGCTTTTCTGAGAGTGTCTCTGATCTCG 2006
Qy CGCAAGTGAACAACAAGTGTCTCAAGAGCGTTCTGGGAGATGAGAGAGCGGTTTACG 2006
Db CGCAAGTGAACAACAAGTGTCTCAAGAGCGTTCTGGGAGATGAGAGAGCGGTTTACG 1620
Qy CGCAAGTGAACAACAAGTGTCTCAAGAGCGTTCTGGGAGATGAGAGAGCGGTTTACG 1561
Db CGCAAGTGAACAACAAGTGTCTCAAGAGCGTTCTGGGAGATGAGAGAGCGGTTTACG 2125
Qy AGTGAATTTGGCTGAGGCTCAAGGAGCGGAGCTGACAGGACACCACTCTGAGCCAGCC 2066

[illegible]

508 GGCATTAATGATAGTATTTAAGTAAAGAGATCCATATGAGTGTGCTTATAT 567
Db 61 GGCATTAATGATAGTATTTAAGTAAAGAGATCCATATGAGTGTGCTTATAT 120
Qy 568 TCACATCTCAAGAGATGATGATTTGTCAATGTCCTTCCCTTTTCTCTCCCT 627
Db 121 TCACATCTCAAGAGATGATGATTTGTCAATGTCCTTCCCTTTTCTCTCCCT 180
Qy 628 CTTTCATCT 687
Db 181 CTTTCATCT 240
Qy 688 CTTTCATCT 747
Db 241 CTTTCATCT 300
Qy 748 AAATTTGACAGATATATACAGAGATGCGCATATACATCTCTATCCACTTTT 807
Db 301 AAATTTGACAGATATATACAGAGATGCGCATATACATCTCTATCCACTTTT 867
Qy 808 TGTGCTATTTGATGCTCAAGTGTGTGACAGAGGCTGACGCGCGAGGCTTCTCAT 420
Db 361 TGTGCTATTTGATGCTCAAGTGTGTGACAGAGGCTGACGCGCGAGGCTTCTCAT 927
Qy 421 GGCATCAACGCTCTACTGAAAGGTGAGACAGAGACAGCTTATCAACCTTCAAGACTG 480
Db 928 GTGTATAGTCAAGTCAAGATGCAAAACAGAGGTATGAAAGGCTTCAAGAGGCTT 987
Qy 481 GTGTATAGTCAAGTCAAGATGCAAAACAGAGGTATGAAAGGCTTCAAGAGGCTT 540
Db 988 CTAACTTAATTTGAGGCGCC - AACAGGCTCCAGAGAAATGCTCACTGAGAGAGACC 1046
Qy 541 CTAACTTAATTTGAGGCGCC - AACAGGCTCCAGAGAAATGCTCACTGAGAGAGACC 600
Db 1047 TGAAGATGAACAGTGGCTAAGCAAGGCTTATATGTGTATTAATGAGTGAATCT 1106
Qy 601 TGAAGATGAACAGTGGCTAAGCAAGGCTTATATGTGTATTAATGAGTGAATCT 660
Db 1107 AATTGGAGAGGAGAGAGGTTGCAAGTGAAGTCAAGGCTTGTGAGCAATGCCAAGG 1166
Qy 661 AATTGGAGAGGAGAGAGGTTGCAAGTGAAGTCAAGGCTTGTGAGCAATGCCAAGG 1226
Db 1167 AATTCTGAAACCTTTAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCA 720
Qy 721 AATTCTGAAACCTTTAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCA 1286
Db 1227 AATTCTGAAACCTTTAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCA 780
Qy 781 AATTCTGAAACCTTTAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCA 840
Db 1287 CGGTCGAAAGCAATGAGCTTTGAAACAGTGTGCAAGGAGCAAGAAAGATTAACA 1346
Qy 841 CGGTCGAAAGCAATGAGCTTTGAAACAGTGTGCAAGGAGCAAGAAAGATTAACA 900
Db 1347 TTTTATTTTAAAGAAATTTGGCCAGGAAATTAATGATTAACGAGAGCAAGAAATTTG 1406
Qy 901 TTTTATTTTAAAGAAATTTGGCCAGGAAATTAATGATTAACGAGAGCAAGAAATTTG 960
Db 1407 GAATGTTAATATAGCTGGGCTATGCGATTTGCTTAAGTGTGCTTTGTTTCTC 1466
Qy 961 GAATGTTAATATAGCTGGGCTATGCGATTTGCTTAAGTGTGCTTTGTTTCTC 1020
Db 1467 TTGAGAAATTAATACTGAGGAGCTTCCCTTTCAAGGCTTATGCGCAACATCTGAC 1526
Qy 1021 TTGAGAAATTAATACTGAGGAGCTTCCCTTTCAAGGCTTATGCGCAACATCTGAC 1586
Db 1527 TTTTCAATATGTTAAGTGTGCTTCCCTTTCAAGGCTTATGCGCAACATCTGAC 1080
Qy 1081 TTTTCAATATGTTAAGTGTGCTTCCCTTTCAAGGCTTATGCGCAACATCTGAC 1140
Db 1587 CAACATGACAGCCAGTCAAAATGCGCCAGAGTCTTTCTGAGTGAATCTCAAGCAATTA 1646

Db 1141 CAACATGACAGCCAGTCAAAATGCGCCAGAGTCTTTCTGAGTGAATCTCAAGCAATTA 1200
Qy 1647 GCGAAGGCTCCCTGATACCGAGAGAGACTGTGCGCTGAGCTGATCTCTTCAAGAC 1706
Db 1201 GCGAAGGCTCCCTGATACCGAGAGAGACTGTGCGCTGAGCTGATCTCTTCAAGAC 1260
Qy 1707 CTCCCAACCTTCCAGAGTGAATCAAGAGAGAACTTTTGAAGAGGAGAGAGAGAGAG 1766
Db 1261 CTCCCAACCTTCCAGAGTGAATCAAGAGAGAACTTTTGAAGAGGAGAGAGAGAGAG 1320
Qy 1767 CTGAGGCTCCCTGATACCGAGAGAGAGTGCAGGCACTGAGAGAGAGAGAGAGAGAG 1826
Db 1321 CTGAGGCTCCCTGATACCGAGAGAGAGTGCAGGCACTGAGAGAGAGAGAGAGAGAG 1380
Qy 1827 TGCGCAATTTGTGCAAGAACTTGAATCTGCTTCAAGCTTCTCTCTCTCTCTCTCT 1886
Db 1381 TGCGCAATTTGTGCAAGAACTTGAATCTGCTTCAAGCTTCTCTCTCTCTCTCTCT 1440
Qy 1887 CGCGGAGAGAGAGAGTGTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1946
Db 1441 CGCGGAGAGAGAGAGTGTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1947 AGGCTTCT 2006
Db 1501 AGGCTTCT 1560
Qy 2007 GCAAGAGTGAACACAGAGTGTCAAAAGAGCTTGTGAGAGAGAGAGAGAGAGAGAG 2066
Db 1561 GCAAGAGTGAACACAGAGTGTCAAAAGAGCTTGTGAGAGAGAGAGAGAGAGAGAG 1620
Qy 2067 GTGACTTGTGAG 2126
Db 1621 GTGACTTGTGAG 1680
Qy 2127 TGAGCTCCAG 2186
Db 1681 TGAGCTCCAG 1727
Qy 2187 TTGGGAG 2246
Db 1728 TTGGGAG 1740
Qy 2247 CTAAGCTTGTGATCCCT 2306
Db 1741 CTAAGCTTGTGATCCCT 1740
Qy 2307 ATGAGCAAGTGGGCGGAG 2366
Db 1741 ATGAGCAAGTGGGCGGAG 1765
Qy 2367 GGCACCTGGAG 2426
Db 1766 GGCACCTGGAG 1765
Qy 2427 AGGACCTTCTCAG 2486
Db 1766 AGGACCTTCTCAG 1765
Qy 2487 GATGCAAGTGTGATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2546
Db 1766 GATGCAAGTGTGATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1765
Qy 2547 TGCGCTCAGCTTATGAGTGGCGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2606
Db 1766 TGCGCTCAGCTTATGAGTGGCGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1767
Qy 2607 GCTTTAGAGTGTGAG 2666
Db 1768 GCTTTAGAGTGTGAG 1827
Qy 2667 CTTGACAG 2726

```

Db      1828 CCTGACGACGCAAGCCAAAGCTCCGCTCCGGCGGGGTGGTGAAGTGGCCGCCGCCGCCGCC 1887
OY      2727 GGGGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2786
OY      1888 GGGGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1947
OY      2787 CAGAGTGAAGAAATTAAGTACAGACAGGACGACGACGACGACGACGACGACGACGACGACGACG 2846
Db      1948 CAGAGCTGAGAAATTAAGTACAGACAGGACGACGACGACGACGACGACGACGACGACGACGACG 2007
OY      2847 GCACTTTTCTTTCGAGAAAAGTTATATGAGGAGCTGAAATGAGCTTCTGAGAGCTTGT 2906
OY      2008 GCACCTTTCTTTCGAGAAAAGTTATATGAGGAGCTGAAATGAGCTTCTGAGAGCTTGT 2067
OY      2907 ACCGTTTTTATTTGTCACACAGAAAAGAACTGCTTCTTCCCTCCGGGAAATCTCT 2966
OY      2068 ACCGTTTTTATTTGTCACACAGAAAAGAACTGCTTCTTCCCTCCGGGAAATCTCT 2127
Db      2967 CTTTAAAGCTGTAAGTGGCTGCGCTGAGTGGTTCAATTTGTTTGTGTTTCTGCGCTTCT 3026
Db      2128 CTTTAAAGCTGTAAGTGGCTGCGCTGAGTGGTTCAATTTGTTTGTGTTTCTGCGCTTCT 2187
OY      3027 CTTTCTCTTTTGGCTTCTTCTTCTTCTGCACTCCCATGCTGATTTCTGCTGCTCTCG 3086
Db      2188 CTTTCTCTTTTGGCTTCTTCTTCTTCTGCACTCCCATGCTGATTTCTGCTGCTCTCG 2247
OY      3087 CTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3146
Db      2248 CTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2307
OY      3147 ACTGAGAGAGCTGTTTGAAGAAGTCCCTGCTCGTCAAGATCCGAGCTTGAGATGGCTAA 3206
Db      2308 ACTGAGAGAGCTGTTTGAAGAAGTCCCTGCTCGTCAAGATCCGAGCTTGAGATGGCTAA 2367
OY      3207 TCAAG 3212
OY      2368 TCAAG 2373
Db      3207 TCAAG 3212
Db      2368 TCAAG 2373

RESULT 8
HSAPT1 1877 bp DNA linear PRI 08-JUN-1995
LOCUS HSAPT1 H.sapiens APT gene, exon 1.
DEFINITION X81335
ACCESSION X81335.1 GI:537410
VERSION X81335.1
KEYWORDS APO-1 gene; apt gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Chordata; Craniata; Vertebrata; Euteleostomi;
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
  1. Behrmann, I., Walczak, H. and Kramer, P. H.
    TITLE Structure of the human APO-1 gene
    JOURNAL Eur. J. Immunol. 24 (12), 3057-3062 (1994)
MEDLINE 95104292
PUBMED 7528667
REFERENCES
  2. (bases 1 to 1877)
    Kramer, P. H.
    Direct Submission
    Submitted (03-SEP-1994) P. H. Kramer, German Cancer Research
    Center, Tumorimmunology Programm, Im Neuenheimer Feld 280, 69120
    Heidelberg, FRG

FEATURES
  source
    1. 1877
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="10"
    /map="10q23"
    /clone="CAPO-1"
    /clone_lib="cosmid, pcos2EMBL"
    151..261
    /note="pyrimidine rich region"

misc_feature
  /note="pyrimidine rich region"

```

Missing Exon 1 and intron

```

gene
  join(1653..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  join(1653..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA1"
  1653..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA1
  /label=ex1a
  join(1670..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA2"
  1670..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA2
  /label=ex1b
  join(1689..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA3"
  1689..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA3
  /label=ex1c
  join(1704..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA4"
  1704..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA4
  /label=ex1d
  join(1707..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA5"
  1707..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA5
  /label=ex1e
  join(1708..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA6"
  1708..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA6
  /label=ex1f
  join(1709..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA7"
  1709..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA7
  /label=ex1g
  join(1710..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA8"
  1710..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA8
  /label=ex1h
  join(1711..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA9"
  1711..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA9
  /label=ex1i
  join(1712..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA10"
  1712..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA10
  /label=ex1j
  join(1713..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA11"
  1713..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA11
  /label=ex1k
  join(1714..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA12"
  1714..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA12
  /label=ex1l
  join(1715..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA13"
  1715..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA13
  /label=ex1m
  join(1716..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA14"
  1716..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA14
  /label=ex1n
  join(1717..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA15"
  1717..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA15
  /label=ex1o
  join(1718..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA16"
  1718..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA16
  /label=ex1p
  join(1719..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA17"
  1719..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA17
  /label=ex1q
  join(1720..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA18"
  1720..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA18
  /label=ex1r
  join(1721..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA19"
  1721..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA19
  /label=ex1s
  join(1722..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA20"
  1722..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA20
  /label=ex1t
  join(1723..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA21"
  1723..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA21
  /label=ex1u
  join(1724..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA22"
  1724..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA22
  /label=ex1v
  join(1725..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA23"
  1725..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA23
  /label=ex1w
  join(1726..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA24"
  1726..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA24
  /label=ex1x
  join(1727..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA25"
  1727..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA25
  /label=ex1y
  join(1728..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA26"
  1728..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA26
  /label=ex1z
  join(1729..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA27"
  1729..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA27
  /label=ex1aa
  join(1730..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA28"
  1730..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA28
  /label=ex1ab
  join(1731..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA29"
  1731..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA29
  /label=ex1ac
  join(1732..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA30"
  1732..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA30
  /label=ex1ad
  join(1733..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA31"
  1733..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA31
  /label=ex1ae
  join(1734..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA32"
  1734..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA32
  /label=ex1af
  join(1735..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA33"
  1735..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA33
  /label=ex1ag
  join(1736..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA34"
  1736..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA34
  /label=ex1ah
  join(1737..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA35"
  1737..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA35
  /label=ex1ai
  join(1738..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA36"
  1738..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA36
  /label=ex1aj
  join(1739..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA37"
  1739..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA37
  /label=ex1ak
  join(1740..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA38"
  1740..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA38
  /label=ex1al
  join(1741..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA39"
  1741..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA39
  /label=ex1am
  join(1742..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA40"
  1742..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA40
  /label=ex1an
  join(1743..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA41"
  1743..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA41
  /label=ex1ao
  join(1744..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA42"
  1744..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA42
  /label=ex1ap
  join(1745..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA43"
  1745..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA43
  /label=ex1aq
  join(1746..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA44"
  1746..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA44
  /label=ex1ar
  join(1747..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA45"
  1747..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA45
  /label=ex1as
  join(1748..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA46"
  1748..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA46
  /label=ex1at
  join(1749..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA47"
  1749..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA47
  /label=ex1au
  join(1750..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA48"
  1750..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA48
  /label=ex1av
  join(1751..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA49"
  1751..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA49
  /label=ex1aw
  join(1752..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA50"
  1752..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA50
  /label=ex1ax
  join(1753..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA51"
  1753..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA51
  /label=ex1ay
  join(1754..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:10
```


Query Match 57.6%; Score 1650.8; DB 9; Length 1877;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1874; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 447 AAGCTTTTGGCTACATTTTATTTTAAAGTAATTTAAATATACATCTACT 506
 DB 1 AAGCTTTTGGCTACATTTTATTTTAAAGTAATTTAAATATACATCTACT 60
 QY 507 GGGCTATATAGATATAGATATAGATATAGATATAGATATAGATATAGAT 566
 DB 61 GGGCTATATAGATATAGATATAGATATAGATATAGATATAGATATAGAT 120
 QY 567 TTGACACTCAAGATATAGATATAGATATAGATATAGATATAGATATAG 626
 DB 121 TTGACACTCAAGATATAGATATAGATATAGATATAGATATAGATATAG 180
 QY 627 TCCCTCATTCCTTCTCCCTTCTCCCTTCTCCCTTCTCCCTTCTCCCT 686
 DB 181 TCCCTCATTCCTTCTCCCTTCTCCCTTCTCCCTTCTCCCTTCTCCCT 240
 QY 687 TCTTTTACATTTTATTTTAAATGAACTTTTCAATTTTGAATAGTTTAA 746
 DB 241 TCTTTTACATTTTATTTTAAATGAACTTTTCAATTTTGAATAGTTTAA 300
 QY 747 AAAATTTGAGAGATATATAGATATAGATATAGATATAGATATAGAT 806
 DB 301 AAAATTTGAGAGATATATAGATATAGATATAGATATAGATATAGAT 360
 QY 807 TTGCTCATTTAGATCTTCAAGATATAGATATAGATATAGATATAGAT 866
 DB 361 TTGCTCATTTAGATCTTCAAGATATAGATATAGATATAGATATAGAT 420
 QY 867 TGGCACTAACAGTCTATGAAAGGTGAAAGGAGGAGGAGGAGGAGGAG 926
 DB 421 TGGCACTAACAGTCTATGAAAGGTGAAAGGAGGAGGAGGAGGAGGAG 479
 QY 927 GGTGTAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 986
 DB 480 GGTGTAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 539
 QY 987 CCTAACCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1045
 DB 540 CCTAACCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 599
 QY 1046 CTGAAGATATAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1105
 DB 600 CTGAAGATATAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 659
 QY 1106 TAATTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1165
 DB 660 TAATTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 719
 QY 1166 GAATCTGAAACCTTTAGTGTGCTGATGATGATGATGATGATGAT 1225
 DB 720 GAATCTGAAACCTTTAGTGTGCTGATGATGATGATGATGATGAT 779
 QY 1226 GATGTCATTTCAAACTACCTTCTGTAATTAATGATGATGATGAT 1285
 DB 780 GATGTCATTTCAAACTACCTTCTGTAATTAATGATGATGATGAT 839
 QY 1286 ACCGTCCTCAAGCAATGATGATGATGATGATGATGATGATGATG 1345
 DB 840 ACCGTCCTCAAGCAATGATGATGATGATGATGATGATGATGATG 899
 QY 1346 ATTTTATTTTAAAGAAATTTGCGAGAAATTAAGATGAGGAGGAGG 1405
 DB 900 ATTTTATTTTAAAGAAATTTGCGAGAAATTAAGATGAGGAGGAGG 959
 QY 1406 TGAATGTTTAAATATAGTGGGCTATGATGATGATGATGATGATG 1465
 DB 960 TGAATGTTTAAATATAGTGGGCTATGATGATGATGATGATGATG 1019
 QY 1466 CTGAGAAATTAATAAATTAAGGAGGAGGAGGAGGAGGAGGAGG 1525

DB 1020 CTGAGAAATTAATAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1079
 QY 1526 CTTTATTAATGATTAATGATGATGATGATGATGATGATGATGATGAT 1585
 DB 1080 CTTTATTAATGATTAATGATGATGATGATGATGATGATGATGATGAT 1139
 QY 1586 ACAACATGAGCAAGCCCAATGATGATGATGATGATGATGATGATG 1645
 DB 1140 ACAACATGAGCAAGCCCAATGATGATGATGATGATGATGATGATG 1199
 QY 1646 AGCCAAAGGCTCTGATGATGATGATGATGATGATGATGATGATG 1705
 DB 1200 AGCCAAAGGCTCTGATGATGATGATGATGATGATGATGATGATG 1259
 QY 1706 CTTTCCCAATTTCCAGATGATGATGATGATGATGATGATGATGAT 1765
 DB 1260 CTTTCCCAATTTCCAGATGATGATGATGATGATGATGATGATGAT 1319
 QY 1766 TCTGAGGCTCTACCTGATGATGATGATGATGATGATGATGATGAT 1825
 DB 1320 TCTGAGGCTCTACCTGATGATGATGATGATGATGATGATGATGAT 1379
 QY 1826 ATGCCATTTGTCAGACCAATGATGATGATGATGATGATGATGAT 1885
 DB 1380 ATGCCATTTGTCAGACCAATGATGATGATGATGATGATGATGAT 1439
 QY 1886 CCGGCGGAGGCAAGGTCGATGATGATGATGATGATGATGATGATG 1945
 DB 1440 CCGGCGGAGGCAAGGTCGATGATGATGATGATGATGATGATGATG 1499
 QY 1946 GAGGCTTCTTCCATCTCTGATGATGATGATGATGATGATGATGAT 2005
 DB 1500 GAGGCTTCTTCCATCTCTGATGATGATGATGATGATGATGATGAT 1559
 QY 2006 CGCAAGATGACACAGGATGATGATGATGATGATGATGATGATGAT 2065
 DB 1560 CGCAAGATGACACAGGATGATGATGATGATGATGATGATGATGAT 1619
 QY 2066 AGTGAATTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2125
 DB 1620 AGTGAATTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1679
 QY 2126 CTGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2185
 DB 1680 CTGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1739
 QY 2186 GTTGGGAAAGCTCTTCACTTCAAGATGATGATGATGATGATGAT 2245
 DB 1740 GTTGGGAAAGCTCTTCACTTCAAGATGATGATGATGATGATGAT 1799
 QY 2246 TCTTCACTTCAAGATGATGATGATGATGATGATGATGATGATGAT 2305
 DB 1800 TCTTCACTTCAAGATGATGATGATGATGATGATGATGATGATGAT 1859
 QY 2306 GATGAGCAAAAGTGGGAGG 2323
 DB 1860 GATGAGCAAAAGTGGGAGG 1877

Mutins Exon 1 and intron

RESULT 9
 HSPFAXL
 LOCUS
 DEFINITION
 H.sapiens Fas, Apo-1 gene (promoter and exon 1).
 ACCESION
 X82279.1 GI:673405
 VERSION
 X82279.1 GI:673405
 KEYWORDS
 APO-1 gene; Fas gene.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 AUTHORS
 Cheng, J., Liu, C., Koopman, W. J. and Mountz, J. D.

TITLE Characterization of human Fas gene. Exon/Intron organization and
 JOURNAL Promoter region
 MEDLINE J. Immunol. 154 (3), 1239-1245 (1995)
 PUBMED 95123075
 REFERENCE 7529798
 AUTHORS 2 (bases 1 to 1608)
 TITLE Cheng J.
 JOURNAL Direct Submission
 COMMENT Submitted (20-OCT-1994) J. Cheng, Univ. of Alabama at Birmingham,
 Birmingham, AL 35294-0007, USA
 Related sequences: M67454 and X63717.
 FEATURES
 source Location/Qualifiers
 1.1608
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="10q23 or 10q24.1"
 /issue_type="placenta"
 /clone_id="Placental genomic DNA"
 /gene="Fas, Apo-1"
 /gene="Fas, Apo-1"
 /gene="Fas, Apo-1"
 BASE COUNT 398 a 421 c 423 g 366 t
 ORIGIN
 Query Match 47.8%; Score 1536.8; DB 9; Length 1608;
 Best Local Similarity 98.6%; Pred. No. 0; Indels 5; Gaps 3;
 Matches 1582; Conservative 0; Mismatches 17;
 Oy 767 AGAGATGCCATATACCATCTCTCTATATCCCATCTCTTTTGTGTCTATTAATGCTCA 826
 Db 7 AAGAGATGCCATATACCATCTCTCTATATCCCATCTCTTTTGTGTCTATTAATGCTCA 66
 Oy 827 GAGTGTGCACAAAGGCTGCGACGCCAGGCTCTTCCATGAGCACTAAAGTCTACTGA 886
 Db 67 GAGTGTGCACAAAGGCTGCGACGCCAGGCTCTTCCATGAGCACTAAAGTCTACTGA 126
 Oy 887 AAGGTGAACAGAGCAAGGCTATCAACCTTCAACAGACTGTGTGATGAGTACAGTACAG 946
 Db 127 AAGGTGAACAGAGCAAGGCTATCAACCTTCAACAGACTGTGTGATGAGTACAGTACAG 186
 Oy 947 ATGCAAAACACAGGCTGATGAAAGCCCTTCAAGAGGCTTACCTAGATTGAGGCG 1006
 Db 187 ATGCAAAACACAGGCTGATGAAAGCCCTTCAAGAGGCTTACCTAGATTGAGGCG 246
 Oy 1007 CC-AACGGCTCCAGAGAAATGTCACTGAGAGAGAGCTGAGATGAAAGTGGCG 1065
 Db 247 CC-AACGGCTCCAGAGAAATGTCACTGAGAGAGAGCTGAGATGAAAGTGGCG 306
 Oy 1066 TAAGCAAGGGTTATTAATGTTTATTAATGCTTGAATCTTAATTGGGAGAGAGAGG 1125
 Db 307 TAAGCAAGGGTTATTAATGTTTATTAATGCTTGAATCTTAATTGGGAGAGAGAGG 366
 Oy 1126 TTGCAAGTGGAGGTCAGAGCTTGTGACATGCAAGAGATCTGAACCTTTAGTG 1185
 Db 367 TTGCAAGTGGAGGTCAGAGCTTGTGACATGCAAGAGATCTGAACCTTTAGTG 426
 Oy 1186 TGTCCAGTGTGAACTGATCAAAATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1245
 Db 427 TGTCCAGTGTGAACTGATCAAAATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 486
 Oy 1246 CCTTCTGTAATTAATTCATGCTAACTAAGAGTATCTAAGCTTCCAAAGCAATAGTG 1305
 Db 487 CCTTCTGTAATTAATTCATGCTAACTAAGAGTATCTAAGCTTCCAAAGCAATAGTG 546
 Oy 1306 ACTTTGAACAGGTTCACAGAGACAGAAAGATTAAGATTTTAAAGAAATTT 1365
 Db 547 ACTTTGAACAGGTTCACAGAGACAGAAAGATTAAGATTTTAAAGAAATTT 606
 Oy 1366 GGCAGGAATTAATGATGATGAGAGAGAGAGATTTGTGAATGTTTAACTAGCTG 1425

Db 607 GGCAGGAATTAATGATGATGAGAGAGAGAGATTTGTGAATGTTTAACTAGCTG 666
 Oy 1426 GGCATATGATTTGGCTTAAGTTTGGTTTCTCTTGAAGATTAATAAATAAG 1485
 Db 667 GGCATATGATTTGGCTTAAGTTTGGTTTCTCTTGAAGATTAATAAATAAG 726
 Oy 1486 GGGCCCTCCCTTTCAAGAGCTTATGAGCAATCTGATCTTTTCAATAGGTTAAG 1545
 Db 727 GGGCCCTCCCTTTCAAGAGCTTATGAGCAATCTGATCTTTTCAATAGGTTAAG 786
 Oy 1546 TCCATTCAGAAAGCTGTGAGCTTCAATGTTGAGAGCAACAATGAGCAAGCCAGTC 1605
 Db 787 TCCATTCAGAAAGCTGTGAGCTTCAATGTTGAGAGCAACAATGAGCAAGCCAGTC 846
 Oy 1606 AAATGCCCGGAGTCTTCTGATGATCTCAAGATTAAGCAAGGCTCTGTAACCA 1665
 Db 847 AAATGCCCGGAGTCTTCTGATGATCTCAAGATTAAGCAAGGCTCTGTAACCA 906
 Oy 1666 GGCAGGACCTGCGCTCTGAGCTTCAATCTCTCAAGACCTCCCACTTCCAGGTT 1725
 Db 907 GGCAGGACCTGCGCTCTGAGCTTCAATCTCTCAAGACCTCCCACTTCCAGGTT 966
 Oy 1726 GAATCAAGCAAGACCTTTGAAAGGAGAGGCGGCTCTGAGGTTCTTCAAGGAA 1785
 Db 967 GAATCAAGCAAGACCTTTGAAAGGAGAGGCGGCTCTGAGGTTCTTCAAGGAA 1845
 Oy 1786 GTGAGCAGTCCAGCACTGACAGAAAGCCCGGAGAGAGATGATGATGAGCAAGAA 1885
 Db 1027 GTGAG-ATGCCAGCCACTGACAGAAAGCCCGGAGAGAGATGATGATGAGCAAGAA 1945
 Oy 1846 CCTGATCTCTTCTGATCCTGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1905
 Db 1086 CCTGATCTCTTCTGATCCTGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1145
 Oy 1906 TGAATCAATGAGAGCTTCCCAAGCCCGGAGAGAGATGATGATGAGCAAGAA 1965
 Db 1146 TGAATCAATGAGAGCTTCCCAAGCCCGGAGAGAGATGATGATGAGCAAGAA 1205
 Oy 1966 CTGACCAAGGAGGCTTCTGAGCTCTGATCTGATCTGATCTGATCTGATCTGATCT 2025
 Db 1206 CTGACCAAGGAGGCTTCTGAGCTCTGATCTGATCTGATCTGATCTGATCTGATCT 1265
 Oy 2026 GTTCAAGAGCTTCTGAGGAGTGAAGAGGAGGTTTACAGATGATCTGATCTGATCT 2085
 Db 1266 GTTCAAGAGCTTCTGAGGAGTGAAGAGGAGGTTTACAGATGATCTGATCTGATCT 1325
 Oy 2086 AGGGGAGGAGCTGAGCAAGAGCAACCTTGAAGGAGGAGGAGGAGGAGGAGGAGG 2145
 Db 1326 AGGGGAGGAGCTGAGCAAGAGCAACCTTGAAGGAGGAGGAGGAGGAGGAGGAGG 1385
 Oy 2146 GCTCTTCTCCGCGGCTTGTGAGCCGCTCAAGTGAAGAGGAGGAGGAGGAGGAGG 2205
 Db 1386 GCTCTTCTCCGCGGCTTGTGAGCCGCTCAAGTGAAGAGGAGGAGGAGGAGGAGG 1445
 Oy 2206 TCGAGAGATTTGATCAACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2265
 Db 1446 TCGAGAGATTTGATCAACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1505
 Oy 2266 CTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2325
 Db 1506 CTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1565
 Oy 2326 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2369
 Db 1566 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1606

RESULT 10
 AX347321/C 1608 bp DNA linear PAT 01-FEB-2002
 LOCUS AX347321
 DEFINITION Sequence 2392 from Patent WO0200928.
 ACCESSION AX347321
 VERSION AX347321.1 GI:18495209

[illegible]

Db 361 GATAGGCAAAAGTGGGGCGGGCGGAGACGCTGCGGATTCGGCGGACGAGGCAAGC 420

QY 2366 GGGGACCTTGGAGCGCGGGCTGCTGGGGAGAGCGCTTGGAGACTGCTCCGGGGGCTGT 2425

Db 421 GGGGACCTTGGAGCGCGGGCTGCTGGGGAGAGCGCTTGGAGACTGCTCCGGGGGCTGT 480

QY 2426 TAGGACCTTCCCTCAGGCGCGGGTCTCAGAACGATGAGAGACTTGTCTTCTTGGGCT 2485

Db 481 TAGGACCTTCCCTCAGGCGCGGGTCTCAGAACGATGAGAGACTTGTCTTCTTGGGCT 540

QY 2486 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGGCTCTCGAGAGCA 2545

Db 541 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGGCTCTCGAGAGCA 600

QY 2546 CTGCGCTCCAGCTTGAAGTGGGCGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGGA 2605

Db 601 CTGCGCTCCAGCTTGAAGTGGGCGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGGA 660

QY 2606 AGCTTTAGGGTCCCTGAGGGGAGACCCCGTTGGAGAGAGA 2647

Db 661 AGCTTTAGGGTCCCTGAGGGGAGACCCCGTTGGAGAGAGA 702

RESULT 14
LOCUS HSA279011 702 bp DNA linear PRI 02-APR-2001
DEFINITION Homo sapiens partial CD95 gene for CD95 antigen (Apo-1 Fas), exon 1.

ACCESSION AJ279011 GI:13539238
VERSION AJ279011.1
KEYWORDS Apo-1 Fas; CD95 antigen; CD95 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Muschen, M., Re, D., Jungnickel, B., Diehl, V., Rajewsky, K. and Kuppers, R.
Somatic mutations of the CD95 gene in human B cells as a side-effect of the germinal center reaction
Unpublished

JOURNAL 2 Muschen, M., Re, D., Brauninger, A., Wolf, J., Hanemann, M.L., Diehl, V., Kuppers, R. and Rajewsky, K.
Somatic mutations of the CD95 gene in Hodgkin- and Reed-Sternberg cells
Unpublished

REFERENCE 3 (bases 1 to 702)
Muschien, M.
TITLE Direct Submission
SUBMITTED (19-SEP-2000) Muschen M., Department of Immunology, Institute for Genetics, LFI B4 R705, Joseph-Stelzmann-Str. 9, 50931 Koeln, GERMANY

COMMENT Related sequences: D31968 X89101 AJ279012 AJ279013.
FEATURES
source location/Qualifiers
1..702
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_type="Reed-Sternberg cells"
/note="Hodgkin's disease tissue"
1..702
/gene="CD95"
41..313
/gene="CD95"
/standard_name="Apo-1 Fas"
/number=1
1..281
/gene="CD95"
/standard_name="Apo-1 Fas"
282..313
/gene="CD95"
/standard_name="Apo-1 Fas"

BASE COUNT 107 a 198 c 264 g 133 t

ORIGIN

Query Match 21.7%; Score 697.2; DB 9; Length 702;
Best local similarity 99.6%; Pred. No. 7,4e-155; Indels 0; Gaps 0;
Matches 699; Conservative 0; Mismatches 3;

1946 GAGGCTTCTCCATCTCTCTGACCAACGAGGCTTTCTGTAGCTGTCTGTACTCTG 2005
1 GAGGCTTCTCCATCTCTCTGACCAACGAGGCTTTCTGTAGCTGTCTGTACTCTG 60

QY 2006 CGCAAGAGTGAACACAGAGTGTCAAGAGCGCTTGGAGAGTGAAGGAGCGGTTACG 2065

Db 61 CGCAAGAGTGAACACAGAGTGTCAAGAGCGCTTGGAGAGTGAAGGAGCGGTTACG 120

QY 2066 AGTGAAGTGGCTGAGAGCTTCAAGAGCGGCGGAGCACTGAGCAACGAGGCGGAGCC 2125

Db 121 AGTGAAGTGGCTGAGAGCTTCAAGAGCGGCGGAGCACTGAGCAACGAGGCGGAGCC 180

QY 2126 CTGGGCTCCAGGCGGAGAGCTGCTTCTCCCGGAGTTGTGACCGGCTCAATACGGA 2185

Db 181 CTGGGCTCCAGGCGGAGAGCTGCTTCTCCCGGAGTTGTGACCGGCTCAATACGGA 240

QY 2186 GTTGGGAGAGCTTCTTCACTTGGAGAGATTTGCTCAACACATGCTGGGAGATTTGAGACC 2245

Db 241 GTTGGGAGAGCTTCTTCACTTGGAGAGATTTGCTCAACACATGCTGGGAGATTTGAGACC 300

QY 2246 TCTTACCTCTGATGATCCCTCTCTGCGGAGGCTTACCCGCTTGTAGTCCCGG 2305

Db 301 TCTTACCTCTGATGATCCCTCTCTGCGGAGGCTTACCCGCTTGTAGTCCCGG 360

QY 2306 GATAGGCAAGTGGGCGGCGGAGCGGAGCGGAGATTTGCGGCGGAGCGGAGCGGAGCGG 2365

Db 361 GATAGGCAAGTGGGCGGCGGAGCGGAGCGGAGATTTGCGGCGGAGCGGAGCGGAGCGG 420

QY 2366 GGGGACCTTGGAGCGGGGCTGCTGGGAGAGGCTTGGAGAGCTGCTCCGGGGCTGT 2425

Db 421 GGGGACCTTGGAGCGGGGCTGCTGGGAGAGGCTTGGAGAGCTGCTCCGGGGCTGT 480

QY 2426 TAGGACCTTCCCTCAGGCGCGGGTCTCAGAACGATGAGAGACTTGTCTTGGGCT 2485

Db 481 TAGGACCTTCCCTCAGGCGCGGGTCTCAGAACGATGAGAGACTTGTCTTGGGCT 540

QY 2486 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGGCTCTCGAGAGCA 2545

Db 541 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGGCTCTCGAGAGCA 600

QY 2546 CTGCGCTCCAGCTTGAAGTGGGCGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGGA 2605

Db 601 CTGCGCTCCAGCTTGAAGTGGGCGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGGA 660

QY 2606 AGCTTTAGGGTCCCTGAGGGGAGACCCCGTTGGAGAGAGA 2647

Db 661 AGCTTTAGGGTCCCTGAGGGGAGACCCCGTTGGAGAGAGA 702

RESULT 15
LOCUS HSA279013 702 bp DNA linear PRI 02-APR-2001
DEFINITION Homo sapiens partial CD95 gene for CD95 antigen (Apo-1 Fas), exon 1, polymorphic allele (+337).

ACCESSION AJ279013 GI:13539242
VERSION AJ279013.1
KEYWORDS Apo-1 Fas; CD95 antigen; CD95 gene.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 Muschen, M., Re, D., Jungnickel, B., Diehl, V., Rajewsky, K. and
 Kuppers, R. Somatic mutations of the CD95 gene in human B cells as a
 side-effect of the germinal center reaction
 JOURNAL Unpublished
 REFERENCE 2 Muschen, M., Re, D., Brauning, A., Wolf, J., Hansmann, M.L., Diehl, V.,
 Kuppers, R. and Rajewsky, K. Somatic mutations of the CD95 gene in Hodgkin- and Reed-Sternberg
 cells
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 702)
 AUTHORS Muschen, M.
 TITLE Direct Submission
 JOURNAL Submitted (19-SEP-2000) Muschen M., Department of Immunology,
 Institute for Genetics, LFI E4 R705, Joseph-Stelzmann-Str. 9, 50931
 Koeln, GERMANY
 COMMENT Related sequences: D31968 X89101 AJ279011 AJ279012.
 FEATURES
 source
 1..702
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /cell_type="Reed-Sternberg cells"
 /note="Hodgkin's disease tissue-polymorphic allele (+337)"
 1..702
 /gene="CD95"
 /standard_name="Apo-1 Fas"
 /number=1
 <1..281
 /gene="CD95"
 /standard_name="Apo-1 Fas"
 282..>313
 /gene="CD95"
 /standard_name="Apo-1 Fas"
 /codon_start=1
 /product="CD95 antigen"
 /protein_id="CAC3541.1"
 /db_xref="GI:13539243"
 /translation="MLGIWTLPLV"
 314..>702
 /gene="CD95"
 /standard_name="Apo-1 Fas"
 /number=1
 BASE COUNT 108 a 198 c 263 g 133 t
 ORIGIN
 Query Match 21.7%; Score 695.6; DB 9; Length 702;
 Best Local Similarity 99.4%; Pred. No. 1.8e-154;
 Matches 698; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1946 GAGGCTTCCTCCATCTCTGACACCGGCGCTTTTCGTGAGCTCGTCTGATCTCG 2005
 DB 1 GAGGCTTCCTCCATCTCTGACACCGGCGCTTTTCGTGAGCTCGTCTGATCTCG 2005
 QY 2006 CGCAGAGTGACACAGGTGTTCAAGAAGCTTCTGGGGAGTGAAGGAAAGCGTTTACG 2065
 DB 61 CGCAGAGTGACACAGGTGTTCAAGAAGCTTCTGGGGAGTGAAGGAAAGCGTTTACG 120
 QY 2066 AGTGACTTGGCTGAGGCTGACGAGGCGCGGCACTGGCAACGAAACACCTGAGGCGCAGCC 2125
 DB 121 AGTGACTTGGCTGAGGCTGACGAGGCGCGGCACTGGCAACGAAACACCTGAGGCGCAGCC 180
 QY 2126 CTGGCTGCCAGGCGGAGCTGCTTCTTCCGCGGGTTGGTGAACCGCTCAATACGGA 2185
 DB 181 CTGGCTGCCAGGCGGAGCTGCTTCTTCCGCGGGTTGGTGAACCGCTCAATACGGA 240

QY 2186 GTTGGGGAAGGCTTTTCACTTCGAGAGATTCTCAACACCATGCTGGGCAATCTGAGACC 2245
 DB 241 GTTGGGGAAGGCTTTTCACTTCGAGAGATTCTCAACACCATGCTGGGCAATCTGAGACC 300
 QY 2246 TCTTAAGCTTGGTATGATCCCTCTCTGCGCGGAGGAGCTTAACCCGCTTTAGTCCGGG 2305
 DB 301 TCTTAAGCTTGGTATGATCCCTCTCTGCGCGGAGGAGCTTAACCCGCTTTAGTCCGGG 360
 QY 2306 GATAGCAAGTGGGAGGCGGCGGCGGAGCCGCTGCGGATTCGCGGCGCAGCGGCGACGC 2365
 DB 361 GATAGCAAGTGGGAGGCGGCGGCGGAGCCGCTGCGGATTCGCGGCGCAGCGGCGACGC 420
 QY 2366 GGGCACTTGGGAGCGGCGGCTGCTGCGGAGGCGGAGTGGAGAGTGGCTCCCGGGGCTGT 2425
 DB 421 GGGCACTTGGGAGCGGCGGCTGCTGCGGAGGCGGAGTGGAGAGTGGCTCCCGGGGCTGT 480
 QY 2426 TAGGACCTTCCTCAAGGCGCGGCTGCTCAAGACGATGAGAGACTTCTTTCTTGGGCTT 2485
 DB 481 TAGGACCTTCCTCAAGGCGCGGCTGCTCAAGACGATGAGAGACTTCTTTCTTGGGCTT 540
 QY 2486 TGATGCAAGTGTCTGATCCGCTGGGCAAGCGGCGAGCTCCGGGCTTCTTGGAGCCA 2545
 DB 541 TGATGCAAGTGTCTGATCCGCTGGGCAAGCGGCGAGCTCCGGGCTTCTTGGAGCCA 600
 QY 2546 CTGCGCTCCACGTTGAGGTGGCGGTGGGGCGGAGCAGGAAATTGAAGCGGAAGTCTGGGA 2605
 DB 601 CTGCGCTCCACGTTGAGGTGGCGGTGGGGCGGAGCAGGAAATTGAAGCGGAAGTCTGGGA 660
 QY 2606 AGCTTTAGGCTGCTGAGAGGGGAGCCCGGTTTGGAGAGAGGA 2647
 DB 661 AGCTTTAGGCTGCTGAGAGGGGAGCCCGGTTTGGAGAGAGGA 702

Search completed: August 1, 2003, 17:32:41
 Job time : 12570.1 secs

Sun Aug 3 09:03:38 2003

US-09-834-291-1.rst

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 10997.1 Seconds
7998.748 Million cell updates/sec

Title: US-09-834-291-1
Perfect score: 3212
Sequence: 1 tgcagctctcagcagatgtaatg.....ttgcagatgcataatcaag 3212

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: em_estbda:*
2: em_estbda:*
3: em_estbda:*
4: em_estbda:*
5: em_estbda:*
6: em_estbda:*
7: em_estbda:*
8: em_estbda:*
9: em_estbda:*
10: em_estbda:*
11: em_estbda:*
12: em_estbda:*
13: em_estbda:*
14: em_estbda:*
15: em_estbda:*
16: em_estbda:*
17: em_estbda:*
18: em_estbda:*
19: em_estbda:*
20: em_estbda:*
21: em_estbda:*
22: em_estbda:*
23: em_estbda:*
24: em_estbda:*
25: em_estbda:*
26: em_estbda:*
27: em_estbda:*
28: em_estbda:*
29: em_estbda:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match Length	ID	Description
No.				
1	710	793	12	BIT63679
2	453.4	14.1	9	AA704610
3	398	12.4	13	BQ072834
4	394.6	12.3	9	AL540709

5	386.4	12.0	506	10	BG659530	BK472931	DKFZP868B
6	381.4	11.9	506	13	BK472931	BF569266	602185634
7	357.8	11.1	599	10	BF569266	BQ924045	AGENCOURT
8	346.2	10.8	1499	13	BQ924045	CB157277	K-EST0216
9	343.6	10.7	524	14	CB157277	BG112070	602352793
10	340.6	10.6	1140	10	BG121070	AV695647	AV695647
11	328.2	10.2	617	9	AV695647	CB124167	K-EST0172
12	306.6	9.5	394	14	CB124167	AA058563	ZE5504.8
13	291.4	9.1	603	9	AA058563	AA056275	ZE53903.8
14	287.4	8.9	577	9	AA056275	AA047220	ZE4941.8
15	280.4	8.7	547	9	AA047220	AA011028	ZE5402.8
16	279.4	8.7	442	9	AA011028	AA057418	ZE59103.8
17	278.4	8.6	447	9	AA057418	AA020992	ZE65107.8
18	276.4	8.6	467	9	AA020992	AL711161	DKFZP686G
19	272	8.5	933	13	AL711161	BO881176	AGENCOURT
20	266.4	8.3	376	14	BO881176	H84298	YB95609.81
21	263	8.2	398	14	H84298	H86126	YB94408.81
22	257	8.0	429	9	H86126	AA239285	XB18D02.Y
23	256.4	8.0	291	14	AA239285	AA018441	ZE50808.8
24	255	7.9	381	9	AA018441	H86544	Y04101.81
25	254.4	7.9	929	14	H86544	BUS53711	AGENCOURT
26	253	7.9	929	13	BUS53711	AL542093	AV692383
27	234.8	7.3	1201	9	AL542093	AV692383	AGENCOURT
28	223.6	7.0	1359	9	AV692383	BG288747	602385566
29	215.4	6.7	359	10	BG288747	AU100132	AGENCOURT
30	214.4	6.7	1339	9	AU100132	BM455788	AGENCOURT
31	210.4	6.6	300	9	BM455788	AL707941	DKFZP686D
32	202.4	6.3	1052	12	AL707941	CD521975	AGENCOURT
33	200.6	6.2	630	9	CD521975	CD521975	AGENCOURT
34	200.2	6.2	741	14	CD521975	CB062973	4013286.B
35	199.8	6.2	775	14	CB062973	BI463384	603204413
36	192.8	6.0	212	14	BI463384	R85827	YQ22910.81
37	191	5.9	800	12	R85827	AV715411	AV715411
38	190	5.9	419	14	AV715411	BI38027	603083875
39	185.2	5.8	696	9	BI38027	CB551158	MSMP0067
40	181.8	5.7	828	12	CB551158	BM922638	AGENCOURT
41	181.4	5.6	600	14	CB551158	BM310113	ih1h11.Y
42	176.4	5.5	1152	12	BM922638	B0662551	cl87f02.Z
43	135	4.2	577	12	BM310113	B0662551	cl87f02.Z
44	131	4.1	658	13	B0662551	AA383161	EST96716
45	124.4	3.9	291	9	AA383161		

ALIGNMENTS

RESULT 1
LOCUS BIT63679/c 793 bp mRNA linear EST 25-SEP-2001
DEFINITION 603049567F1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5189752 5',
mRNA sequence.
ACCESSION BIT63679 GI:15755257
VERSION BIT63679.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC http://mgi.cni.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gga@nsl.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:
http://image.llnl.gov
Plate: L14M11474 Row: e Column: 17
High quality sequence stop: 786.

Query Match	22.18;	Score 710;	DB 12;	Length 793;
Best Local Similarity	97.48;	Pred. No.	1e-111;	
Matches 775;	Conservative	0.	Winn	

2381 GCGGGCTGCTCGGGAGAGCCTTTGAGACTGGCTCCCGGGGGCTGTATAGACTCTTCCCTCA 677
 676 CAGGGCTGCTGCGGGAGCGCTTTGGAACTGCGCTCCCGGGGCTGTATAGACTCTTCCCTCA 2440
 2441 GGGCCGGGCTCTCAGAACATGAGAGACTGTCTTTCTTTGGGCTTATGCGAAGTCTG 617
 616 GGCCCGGGGCTCAGAACCTCTGAGAGACTTCTTTCTTTGGGCTTATGCGAAGTCTG 2500
 2501 ATCCCGCTGGGAGAGCGGGGAGAGCTCCGGGCTCTCTCGAAGCACTGAGGCTTAACTTTC 557

[illegible][illegible]

316 ACCCTGCAGCCTTCGAAACAGTATGTCTATTTCTGCAGGTTCTCAGACGAGGAAT 2800
2801 AAGTCAGACCGGAAGCAGCTGGTTAAGCCGAGGGCTCGAAGAAACGACACTTTTCTTTC 257
256 AAGTCAGACCGGAAGCAGTGGTTAAGCCGAGGGCTCGAAGAAACGACACTTTTCTTTC 2860
2861 TCGAAAG 197

196 TCGAAAAGTTAATATGGGGCTGAAGACCTTGTGAGCGTTGTTACCGTTT-TTTATT 2919
2920 GTCAACAGAAAAGAAAATCGCTTGTCTCCCTTCGGGAAATCTCTCTTTAAACCTGTA 137
136 GTCAACAGAAAAGAAAATCGCTTGTCTCCCTTCGGGAAATCTCTCTTTAAACCTGTA 2979

RESULT 2	
AA704610	
LOCUS	
DEFINITION	AA704610 zj19h02.s1 Soares fetal liver_spleen_infls_s1 Homo sapiens cDNA
ACCESSION	AA704610
VERSION	AA704610.1 GI:2714528
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 46);
Haller L., Allen M., Bowles, L., Dubuque, T., Geisel, G., Joet, S.,
Kizman, D., Kucaba, J., Moore, B., Schell, M., Le, N., Lennon, G.,
White, Y., Wylie, T., Macerston, R., Stepien, M., Tan, F.,
Washu-Nzi human EST project
Unpublished
Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estowatson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGS Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. EF from Amersham
High quality sequence stop: 459.
Location/Qualifiers
1..467
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1387127"
/db_xref="t3aom:9606"
/cclone="IMAGE:450771"
/sex="male"
/dev_stage="20 week-post conception embryo"

```

```

/seq_mofc="DH10B (ampicillin resistant)"
/clonelib="Soares fetal liver_cscn_INFS.S1"
/note="Organ: Liver and Spleen; Vector: pT7SD (Pharmacia)
This is a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
AATCGAGAGATTTATTATTC

```

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT773 vector. Library went through one round of normalisation. Library constructed by Bento Soares and M. Patricia Bonaldo."

```

Query Match      14.1%;   Score 453.4;   DB 9;   Length 467;
Best Local Similarity 99.6%;
Matches 467;   Conservative 0;   Mismatches 1;   Indels 1;   Gaps 1,
694  ACATTTTATTTAATGAACCTTTTCAATTTGGAATAGATTAGATTCAAAAATTT
      |
      |
      |
1  ACATTTTATTTAATGAACCTTTTCAATTTGGAATAGATTAGATTCAAAAATTT 753

```

754 GCAGAGATAATACAGAGAGATGCCATATACATCCCTTATCCCACTTTTGTGTC 60

Sun Aug 3 09:03:38 2003

UB-09-834-291-1.rst

Page 3

DB 61 GCACAGATATACAGAGATGATCCATATACATCCCTTATCCACCTCTTTGTGTC 120
DB 814 TATTAGATGCTCAGAGTGTGCAAGAGCTGCGACGCCAGGCTCTTCCATGAGCACT 873
DB 121 TATTAGATGCTCAGAGTGTGCAAGAGCTGCGACGCCAGGCTCTTCCATGAGCACT 180
DB 874 AACAGTCTACTGAAAGGTGAAACAGAGACAGAGCTTATCAACACTCAAGACTGTGTA 933
DB 181 AACAGTCTACTGAAAGGTGAAACAGAGACAGAGCTTATCAACACTCAAGACTGTGTA 240
DB 934 AGTGCAGTGCAGATGCAAAACAGAGGTGATGAAAGCCCTCAAGAGGTTAACTTACC 993
DB 241 AGTGCAGTGCAGATGCAAAACAGAGGTGATGAAAGCCCTCAAGAGGTTAACTTACC 300
DB 994 TAGATTGAGGGGCC-ACAGGCTCCAGAAAGAAATGTCACTGAGAGAAAGCTGAAAG 1052
DB 301 TAGATTGAGGGGCC-ACAGGCTCCAGAAAGAAATGTCACTGAGAGAAAGCTGAAAG 360
DB 1053 ATGAAAGTGGGCTAGCAAGAGGTTATTAATGCTTATTAATGAGGTTGAATCTAATGG 1112
DB 361 ATGAAAGTGGGCTAGCAAGAGGTTATTAATGCTTATTAATGAGGTTGAATCTAATGG 420
DB 1113 GAAGGAGAGAGGTTGCAAGAGTGAAGTGAAGCTTGTGAGAGT 1159
DB 421 GAAGGAGAGAGGTTGCAAGAGTGAAGTGAAGCTTGTGAGAGT 467

RESULT 3 1089 bp mRNA linear EST 02-APR-2002
LOCUS B0072834
DEFINITION AGENCOURT_6763280 NIH_MGC_118 Homo sapiens cDNA IMAGE:5756324
5' mRNA sequence.

ACCESSION B0072834
VERSION B0072834.1 GI:19901880
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1089)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNL12796, row: 1 column: 21
High quality sequence stop: 645.
Location/Qualifiers
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5756324"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is destroyed upon and directionally cloned (EcoRV site is destroyed for cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

FEATURES

source
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5756324"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is destroyed upon and directionally cloned (EcoRV site is destroyed for cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

BASE COUNT 280 a 263 c 333 g 212 t 1 others
ORIGIN
Query Match 12.4%; Score 398; DB 13; Length 1089;
Best Local Similarity 98.8%; Pred. No. 3.1e-58;
Matches 401; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 2310 GGCAGAGTGGGCGGGGCGGGGAGCGCTGCGGAGTTGCGGCGGAGCGGCGGCGGCGG 79
DB 20 GGCAGAGTGGGCGGGGCGGGGAGCGCTGCGGAGTTGCGGCGGAGCGGCGGCGGCGG 2429
DB 2370 ACTTGGGAGCGGCGGCGGCTGCGGAGCGGCTTGAAGACTGCTCCCGGGGCTTTAAG 139
DB 80 ACCTGGAGCGGCGGCGGCTGCGGAGCGGCTTGAAGACTGCTCCCGGGGCTTTAAG 2489
DB 2430 ACCTTCCCTCAGGCGGCGGCTGCGGAGCGGCTTGAAGACTGCTCCCGGGGCTTTAAG 199
DB 140 ACCTTCCCTCAGGCGGCGGCTGCGGAGCGGCTTGAAGACTGCTCCCGGGGCTTTAAG 2549
DB 2490 GCGAAGTGGTATCCGCTGCGGAGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 259
DB 200 GCGAAGTGGTATCCGCTGCGGAGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 2609
DB 2550 GCTCCAGTGGAGTGGGCGGCTGCGGAGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 2669
DB 260 GCTCCAGTGGAGTGGGCGGCTGCGGAGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 319
DB 2610 TTAGGCTGCTGGAGGAGGAGCGGCTGCGGAGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 379
DB 320 TTAGGCTGCTGGAGGAGGAGCGGCTGCGGAGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 2715
DB 2670 GACAGCCAGCCAGCCAGGCTCCGCTCCGCGGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 425
DB 380 GACAGCCAGCCAGCCAGGCTCCGCTCCGCGGCGGCGGAGCTCCGCGCTCTTGGAGACCACTGC 425

RESULT 4 962 bp mRNA linear EST 12-MAY-2003
LOCUS AL540709/c
DEFINITION AL540709 Homo sapiens PLACENTA Homo sapiens cDNA CSDB002VN18
5-PRIME, mRNA sequence.

ACCESSION AL540709
VERSION AL540709.2 GI:30544172
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 962)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL On Feb 15, 2001 this sequence version replaced gi:12871113.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqef@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5554.r for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgt-bin/cluster.cgi?seq=CSDB002D009QPLcluster=5554.r. Contact :
Feng Liang Email: fliang@life.techn.com URL: <http://life.techn.com/Invitrogen>
http://filllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CSDB002D009QPL.

FEATURES

source
1..962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDB002VN18"
/tissue_type="PLACENTA"
/clone_1ib="Homo sapiens PLACENTA"

us-09-834-291-1.rst

1

Query Match 10.7%; Score 343.6; DB 14; Length 524;
 Best Local Similarity 98.9%; Pred. No. 7.2e-49;
 Matches 346; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

2289 CCCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 2348
 356 CCCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 297
 2349 GCGGCAAGCGGCGCAAGCGGGGCACTTGGAGCGCGGGGCTGTGGAGAGCGCTTGAAGAC 2408
 296 GCGGCAAGCGGCGCAAGCGGGGCACTTGGAGCGCGGGGCTGTGGAGAGCGCTTGAAGAC 237
 2409 TGGCTCCCGGGGGCTGTAGAGACCTTCCCTCAGGCGCGGGGCTGTAGAGAGAGAC 2468
 236 TGGCTCCCGGGGGCTGTAGAGACCTTCCCTCAGGCGCGGGGCTGTAGAGAGAGAC 177
 2469 TTGCTTTCTTGGGCTTGTAGAGAGTGTGATCCCGCTGGGCAAGCGGGGCAAGCTCCG 2528
 176 TTGCTTTCTTGGGCTTGTAGAGAGTGTGATCCCGCTGGGCAAGCGGGGCAAGCTCCG 117
 2529 GCGCTCCCTCGGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGAGAGAAAT 2588
 116 GCGCTCCCTCGGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGAGAGAAAT 57
 2589 GAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGGAGGGGAGACCCCGGTTG 2638
 56 GAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGGAGGGGAGACCCCGGTTG 7

RESULT 10
 BG121070/c 1140 bp mRNA linear EST 30-JAN-2001
 LOCUS 602352793p1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4451039 5',
 DEFINITION mRNA sequence.

ACCESSION BG121070
 VERSION BG121070.1 GI:12614579
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLML at:
 http://image.llnl.gov/
 Plate: L1AM10237 row: 1 column: 24
 High quality sequence start: 9
 High quality sequence stop: 727.
 Location/Qualifiers

FEATURES
 source 1..1140
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4451039"
 /issue_type="adenocarcinoma, cell line"
 /lab_host="PHIOB (phage-resistant)"
 /clone_id="NIH_MGC_90"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."

BASE COUNT 289 a 342 c 331 g 178 t

ORIGIN

Query Match 10.6%; Score 340.6; DB 10; Length 1140;
 Best Local Similarity 98.4%; Pred. No. 2.2e-48;
 Matches 365; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

2289 CCCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 2348
 375 CCCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 316
 2349 GCGGCAAGCGGCGCAAGCGGGGCACTTGGAGCGCGGGGCTGTGGAGAGCGCTTGAAGAC 2408
 315 GCGGCAAGCGGCGCAAGCGGGGCACTTGGAGCGCGGGGCTGTGGAGAGCGCTTGAAGAC 256
 2409 TGGCTCCCGGGGGCTGTAGAGACCTTCCCTCAGGCGCGGGGCTGTAGAGAGAGAC 2468
 255 TGGCTCCCGGGGGCTGTAGAGACCTTCCCTCAGGCGCGGGGCTGTAGAGAGAGAC 196
 2469 TTGCTTTCTTGGGCTTGTAGAGAGTGTGATCCCGCTGGGCAAGCGGGGCAAGCTCCG 2527
 195 TTGCTTTCTTGGGCTTGTAGAGAGTGTGATCCCGCTGGGCAAGCGGGGCAAGCTCCG 136
 2528 GCGCTCCCTCGGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGAGAGAAAT 2587
 135 GCGCTCCCTCGGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGAGAGAAAT 76
 2588 TGAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGGAGGGGAGACCCCGGTTGAGAGAG 2646
 75 TGAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGGAGGGGAGACCCCGGTTGAGAGAG 16
 2647 AGCGGAAGCTCC 2657
 15 AGCGGAAGCTTC 5

RESULT 11
 AV695647/c 617 bp mRNA linear EST 16-JAN-2002
 LOCUS AV695647 GKC Homo sapiens cDNA clone GKCPCD11 5', mRNA sequence.

ACCESSION AV695647
 VERSION AV695647.1 GI:10297510
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLML at:
 http://image.llnl.gov/
 Plate: L1AM10237 row: 1 column: 24
 High quality sequence start: 9
 High quality sequence stop: 727.
 Location/Qualifiers

FEATURES
 source 1..1140

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GKCPCD11"
 /issue_type="hepatocellular carcinoma"
 /dev_stage="Adult"

This clone is available at CHGC in Shanghai.
 Email: hanzg@chgc.sh.cn
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Contact: Zeguangu Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

/lab host="SOLR"
/clone.lib="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 122 a 210 c 179 g 104 t 2 others
ORIGIN

Query Match 10.2%; Score 328.2; DB 9; Length 617;
Best Local Similarity 98.0%; Pred. No. 3.1e-46;
Matches 344; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

2289 CCGCTTACTTCCCGGGATGACCAAGTGGGCGGGCGGAGCGGATGGGATTTGCG 2348
347 CCGCTTACTTCCCGGGATGACCAAGTGGGCGGGCGGAGCGGATTTGCG 288
2349 GCGGACGCGGCGGACCGGCGGACCTGGGAGCGGCGGCTGCTGGGAGCGTTGAGAC 2408
287 GCGGACGCGGCGGACCGGCGGACCTGGGAGCGGCGGCTGCTGGGAGCGTTGAGAC 228
2409 TGGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 2468
227 TGGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 168
2469 TGGGCTTCTTGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 167
167 TTGC-----CTTGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 112
2529 GCGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 2588
111 GCGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 52
2589 GAAAGCGGAAGTCTGGAAGCTTTAGAGGCTGCTGAGAGGCGGACCCCGGTTGG 2639
51 GAAAGCGGAAGTCTGGAAGCTTTAGAGGCTGCTGAGAGGCGGACCCCGGTTGG 1

RESULT 12
LOCUS CBI24167/c 394 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST012829 LSHLK1s1 Homo sapiens cDNA clone LSHLK1s1-11-C12 5',
mRNA sequence.
ACCESSION CBI24167
VERSION CBI24167.1 GI:28084072
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Kim,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

21C Frontier Korean EST Project 2001
TITLE Unpublished
JOURNAL Contact: Kim YS
COMMENT Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 11 row: C column: 12
High quality sequence stop: 394.
Location/Qualifiers
1. 394

FEATURES
Source
1. 394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LSHLK1s1-11-C12"
/sex="M"
/cell_line="HLK-1"
/lab_host="Top10P"
/clone.lib="LSHLK1s1"

/note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
Site_2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then dephosphorylated
with tobacco acid pyrophosphatase (TAP). The dephosphorylated
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 - 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10P with
electroporation method."

BASE COUNT 75 a 149 c 109 g 61 t
ORIGIN

Query Match 9.5%; Score 306.6; DB 14; Length 394;
Best Local Similarity 98.7%; Pred. No. 1.1e-42;
Matches 309; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

2289 CCGCTTACTTCCCGGGATGACCAAGTGGGCGGGCGGAGCGGATTTGCG 2348
315 CCGCTTACTTCCCGGGATGACCAAGTGGGCGGGCGGAGCGGATTTGCG 256
2349 GCGGACGCGGCGGACCGGCGGACCTGGGAGCGGCGGCTGCTGGGAGCGTTGAGAC 2408
255 GCGGACGCGGCGGACCGGCGGACCTGGGAGCGGCGGCTGCTGGGAGCGTTGAGAC 196
2409 TGGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 2468
195 TGGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 136
2469 TGGGCTTCTTGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 2528
135 TGGGCTTCTTGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 76
2529 GCGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 2588
75 GCGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGCGGCTGCTGAGAACATGAGAC 16
2589 GAAAGCGGAAGTCT 2601
15 GAAAGCGGAAGTCT 3

RESULT 13
LOCUS AA058563
DEFINITION z155b04.s1 Soares retina N2bHR Homo sapiens cDNA clone
IMAGE:380815 3', mRNA sequence.
ACCESSION AA058563
VERSION AA058563.1 GI:1551370
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 603)
 AUTHORS Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chissee, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hardis, E., Moore, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Parsons, J., R., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, U., Trevaekis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags

TITLE Genome Res. 6 (9), 807-828 (1996)
 JOURNAL 97044478
 MEDLINE 8885549
 PUBMED

COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 921 Std Error: 0.00
 Seq primer: -40M13 fwd. from Amersham
 High quality sequence stop: 467.
 Location/Qualifiers

FEATURES
 source 1..603
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1289072"
 /db_xref="taxon:9606"
 /clone="IMAGE:380815"
 /sex="male"
 /tissue_type="retina"
 /dev_stage="55 year old"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares retina N2b4R"
 /note="Organ: eye; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A) + RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 145 a 142 c 140 g 175 t 1 others
 ORIGIN

Query Match 9.1%; Score 291.4; DB 9; Length 603;
 Best Local Similarity 99.7%; Pred. No. 6.4e-40;
 Matches 292; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2920 GTGACACAGAAAGAAAGAAAGCTGCTGCTCCGCGGAATCTCTTTAAGCTGTA 2979
 1 GTACACAGAAAGAAAGAAAGCTGCTGCTCCGCGGAATCTCTTTAAGCTGTA 60
 2980 AGTGGCTGCTGAGTGTTCATTTGTTTCTGCGCTCTCTTTCTTTG 3039
 61 AGTGGCTGCTGAGTGTTCATTTGTTTCTGCGCTCTCTTTCTTTG 120
 3040 CCCTTTCTAGCTGCACTCCCATGATGATTTGCTGCTGCTGCGGGGTGTG 3099
 121 CCCTTTCTAGCTGCACTCCCATGATGATTTGCTGCTGCTGCGGGGTGTG 180
 3100 TACTGCTTCCACCGCAGAAAGCCGGCCATATATGGCAAGAACTTGACACCT 3159
 181 TACTGCTTCCACCGCAGAAAGCCGGCCATATATGGCAAGAACTTGACACCT 240
 3160 GTTTGAAAAGTCCCTGCTGCAAGATGCACTTGCAATGCAAG 3212
 241 GTTTGAAAAGTCCCTGCTGCAAGATGCACTTGCAATGCAAG 293

RESULT 14
 AA056275 577 bp mRNA linear EST 17-SEP-1996
 LOCUS 2153903.81 Soares retina N2b4R Homo sapiens cDNA clone
 DEFINITION IMAGE:380692.3, mRNA sequence.

ACCESSION AA056275
 VERSION AA056275.1 GI:1548679
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 577)
 AUTHORS Hallier, L., Clark, N., Dubuque, T., Ellston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaekis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R. The Mashu-Merck EST Project

TITLE Unpublished
 JOURNAL
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -40M13 fwd. from Amersham
 High quality sequence stop: 178.
 Location/Qualifiers

FEATURES
 source 1..577
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1288949"
 /db_xref="taxon:9606"
 /clone="IMAGE:380692"
 /sex="male"
 /tissue_type="retina"
 /dev_stage="55 year old"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares retina N2b4R"
 /note="Organ: eye; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A) + RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 133 a 135 c 135 g 163 t 11 others
 ORIGIN

Query Match 8.9%; Score 287.4; DB 9; Length 577;
 Best Local Similarity 98.3%; Pred. No. 3.1e-39;
 Matches 288; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

2920 GTGACACAGAAAGAAAGAAAGCTGCTGCTCCGCGGAATCTCTTTAAGCTGTA 2979
 1 GTACACAGAAAGAAAGAAAGCTGCTGCTCCGCGGAATCTCTTTAAGCTGTA 60
 2980 AGTGGCTGCTGAGTGTTCATTTGTTTCTGCGCTCTCTTTCTTTG 3039
 61 AGTGGCTGCTGAGTGTTCATTTGTTTCTGCGCTCTCTTTCTTTG 120
 3040 CCCTTTCTAGCTGCACTCCCATGATGATTTGCTGCTGCTGCGGGGTGTG 3099

BASE COUNTY
ORIGIN

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1286557"
/db_xref="taxon:9606"
/clone="IMAGE:380300"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="RDH10B (ampicillin resistant)"
/clone_lib="Soares retina N2bHR"
/note="Organ: eye; Vector: pRT73D (Pharmacia) with a
strand cDNA was primed with a Not I - Site 2; Eco RI; 1st
tGTTACCAACTCTAGAGGAGGAGCGCGCCCTTTTCTTTTCTTTT 3';
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and Eco RI
the Not I and Eco RI sites of a modified pRT73 vector
(Pharmacia). The retinas were digested with Not I and
Caucasian and total cellular poly(A)+ RNA was extracted
hrs after their removal. The retinal RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento
Soares and M.Felicia Bonaldo."

```

```

Query Match      8.7%; Score 280.4; DB 9; Length 547;
Best Local Similarity 99.3%; Pred. No. 5e-38;
Matches 292; Conservative 0; Mismatches 1; Indels 1; Gaps 1.
QY      2920 GTGACACAGAAAAGAAACGCGCTGTCTCCCTTCGCGAAATTCCTCTTAAGACGTGA 2979
Db      1      GTACACACAGAAAAGAAACGCGCTGTCTCCCTTCGCGAAATTCCTCTTAAGACGTGA 2979
QY      2980 AGTCGCGCTAGAGCGTTTCAATTTGTTTGTATTTTGTGCGCTTCCTCTTCTCTTTG 3039
Db      61      AGTCGCTGCTTAGAGCGTTTCAATTTGTTTGTATTTTGTGCGCTTCCTCTTCTCTTTG 3039
QY      3040 CCCTTTCTTAAGCTTGACCTCCACATGAGTGATTTGCTGCTGTCTCTGCGGGTGGTGG 3099
Db      121     CCCTTTCTTAAGCTTGACCTCCACATGAGTATTTCTGCTGTCTCTGCGGGTGGTGG 3099
QY      3100 TACTGCTGCCACGCGACAGAACCGCGCGCTATATATGCGCAGAAACTTGAAGACGCT 3159
Db      181     TACTGCTGCCACGCGACAGAACCGCGCGCTATATATGCGCAGAAACTTGAAGACGCT 3159
QY      3160 GTTTTGAAGATCCCTTCGCTCAGAAATGCGAGCTTGC-AGATGCGCTAATCAAG 3212
Db      241     GTTTTGAAGATCCCTTCGCTCAGAAATGCGAGCTTGCAGAAATGCGCTAATCAAG 294

```


Sun Aug 3 09:03:34 2003

US-09-834-291-1.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Inc.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 1075.81 Seconds
(without alignments)
8059.612 Million cell updates/secTitle: US-09-834-291-1
Perfect score: 3212
Sequence: 1 tggagagcttcaggaatg.....ttgcagatggtatcaaaag 3212Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: N Geneseq.19jun03:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2143.4	66.7	24	ABN79677
2	1536.8	47.8	17	ABT34162
3	961	29.9	16	ABJ34419
4	957.4	29.8	24	ABJ34418
5	287.4	8.9	21	AAJ38177
6	262.8	8.2	26	AAJ28700
7	252	7.8	21	AAJ28700
8	252	7.8	24	ABV94152

9	252	7.8	2551	24	ABN79588
10	252	7.8	2551	25	ABX10416
11	226	7.0	1024	25	AB283782
12	226	7.0	2471	16	AAQ3879
13	226	7.0	2471	20	AAQ24878
14	226	7.0	2534	13	AAQ29959
15	226	7.0	2534	16	AAQ5297
16	226	7.0	2534	17	AAJ16303
17	226	7.0	2534	19	AAJ29993
18	224.4	7.0	2534	18	AAV07002
19	224.4	1.9	260209	24	ABJ55564
20	60	1.9	2435	22	AAJ68654
21	60	1.9	2435	25	ABK3765
22	60	1.9	2435	25	ACA11094
23	60	1.9	2821	24	ACA02280
24	60	1.9	2821	24	ABK39630
25	60	1.9	2821	25	ACA11959
26	60	1.9	2821	25	ACA03145
27	59.8	1.9	35100	19	AAV20441
28	59.8	1.9	38258	20	AAJ28653
29	59.2	1.8	8888	20	AAK02665
30	58.6	1.8	77425	24	ABK3502
31	58.4	1.8	9947	24	ABK52825
32	58	1.8	43599	24	ABK84242
33	57.2	1.8	1000	21	AAA02484
34	57	1.8	857	17	AAJ4530
35	57	1.8	920	17	AAJ4529
36	57	1.8	920	17	ABK64693
37	57	1.8	975	17	AAJ34528
38	57	1.8	975	24	ABJ63728
39	57	1.8	975	24	ABJ68887
40	57	1.8	1104	17	AAJ34527
41	57	1.8	1167	17	AAJ34526
42	56.8	1.8	62488	24	AAJ44981
43	56.4	1.8	2502	25	ABJ76239
44	56.2	1.7	36901	20	AAJ23892
45	56.2	1.7	38886	20	AAJ23897

ALIGNMENTS

RESULT 1
ID ABN79677 standard; DNA; 2165 BP.

AC ABN79677; (first entry)

DT 29-JUL-2002 (first entry)

DE Sequence #1 used to generate target oligonucleotides.

XX Human; immunosuppressive; antiinflammatory; hepatotropic;

KW Cytostatic; vasotropic; hepatitis; cancer; allograft rejection;

KW de; Fas.

XX Homo sapiens.

OS Homo sapiens.

XX Key 1782..1813

XX CDS /product= "peptide encoded by sequence
used create target oligonucleotides"

Human Fas locus HS
DNA encoding prote
Toxicologically re
Fas-delet-TM CDNA.
Soluble Fas recep
Human cell surface
Plasmod pf58 contg
hFas coding sequen
Fas CDNA. Mammali
Human Fas antigen
Human Fas antigen
Human SUT2 gene
Human lung tumour
CDNA encoding L801
Human lung cancer-
lung cancer therap
CDNA encoding L801
Human lung cancer-
lung cancer therap
Human c-fms oncoge
Human c-fms oncoge
Human c-fms oncoge
EP-892047 Seq ID 4
Human CDNA differe
Genomic DNA encodi
Human colon cancer
Human Fas soluble
Human Fas soluble
Human CDNA differe
Human Fas soluble
Human Fas soluble
Kidney cancer rela
Breast cancer rela
Human Fas antige
Human ERM1 splice
Lung cancer-associ
Murine LOBO genom
Murine LOBO homolo


```

Db      67 GAGTGTGTGACAAAGGCTGCGACGCCAGGGTCTTCTCAATGCGACATACTCACTGA 126
QY      887 AAGGTGGAACAGAGACAGGCTATCAACACTTCAAGACTGTGTGTAGTGAAGTGAAG 946
Db      127 AAGGTGGAACAGAGACAGGCTATCAACACTTCAAGACTGTGTGTAGTGAAGTGAAG 946
QY      947 ATGCAAAACACAGGCTGTATGTAAGGCTCTCAGAGAGGTACCTTAACCTTAATTTGAGG 186
Db      187 ATGCAAAACACAGGCTGTATGTAAGGCTCTCAGAGAGGTACCTTAACCTTAATTTGAGG 1006
QY      1007 CC-AACAGGCTCCAGAAAGAAATGTCACTGAGAGAGGCTGAGAGTGAACAGTGGGC 246
Db      247 CCAGACAGGCTCCAGAAAGAAATGTCACTGAGAGAGGCTGAGAGTGAACAGTGGGC 306
QY      1066 TAAAGCAAGGCTTATTAATGTATTAATGAGGCTTAACTTAATTTGGAGAGAGAGG 1125
Db      307 TAAAGCAAGGCTTATTAATGTATTAATGAGGCTTAACTTAATTTGGAGAGAGAGG 366
QY      1126 TTGCAAGTGAAGTGCAGAGCTGTGTGAGAGATGCCAAGAACTAAGAACTTTAGTG 1185
Db      367 TTGCAAGTGAAGTGCAGAGCTGTGTGAGAGATGCCAAGAACTAAGAACTTTAGTG 1185
QY      1186 TGTCCAGTCTGGAACCTGCAATCAATTCAGGTTCAAGTAAGTGTCAATTCAGAACTA 1245
Db      427 TGTCCAGTCTGGAACCTGCAATTCAGGTTCAAGTAAGTGTCAATTCAGAACTA 1245
QY      1246 CCTTCGTAAATTTATGCTAATCACTAAGAGCTTCAACGTTCCAAAGCAATAGTG 1305
Db      487 CCTTCGTAAATTTATGCTAATCACTAAGAGCTTCAACGTTCCAAAGCAATAGTG 546
QY      1306 ACTTTGAACAGTGTTCACAGAGACAGAAAGAAATTAACAAGATTTTAAAGAAAT 1365
Db      547 ACTTTGAACAGTGTTCACAGAGACAGAAAGAAATTAACAAGATTTTAAAGAAAT 606
QY      1366 GGCAGAGAAATTAATGAGTAACAGAGACAGAAAGTAATTTGAATTTAATAGCTG 1425
Db      607 GGCAGAGAAATTAATGAGTAACAGAGACAGAAAGTAATTTGAATTTAATAGCTG 666
QY      1426 GGCAGAGAAATTTGCTTAAGTGTAGCTTTGCTTCTGAGAAATTAATAGCTG 1485
Db      667 GGCAGAGAAATTTGCTTAAGTGTAGCTTTGCTTCTGAGAAATTAATAGCTG 726
QY      1486 GGCAGAGAAATTTGCTTAAGTGTAGCTTTGCTTCTGAGAAATTAATAGCTG 1545
Db      727 GGCAGAGAAATTTGCTTAAGTGTAGCTTTGCTTCTGAGAAATTAATAGCTG 786
QY      1546 TCCATTCCAGAAAGCTGTGAGCTCTCATGTGAGAGCAACATGAGAGAGCTG 1605
Db      787 TCCATTCCAGAAAGCTGTGAGCTCTCATGTGAGAGCAACATGAGAGAGCTG 846
QY      1606 AATATGCCCCGCAAGTCTTCTGAGTGAATTCAGCAATTAAGCCAGAGCTCTGTAACA 1665
Db      847 AATATGCCCCGCAAGTCTTCTGAGTGAATTCAGCAATTAAGCCAGAGCTCTGTAACA 1725
QY      1666 GGCAGAGAACTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 906
Db      907 GGCAGAGAACTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 966
QY      1726 GAACTAAGCAAGAGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1785
Db      967 GAACTAAGCAAGAGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1845
QY      1786 GTGAGCAATGCAAGCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1905
Db      1027 GTGAGCAATGCAAGCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1965
QY      1846 CCTTGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1965
Db      1086 CCTTGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      1906 TGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```

```

Db      1146 TGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      1966 CTGACCAACCGGAGGCTTTTGTGAGCTGTCTGATCTTGGCAAGAGTGAACAAGCT 1205
Db      1206 CTGACCAACCGGAGGCTTTTGTGAGCTGTCTGATCTTGGCAAGAGTGAACAAGCT 2025
QY      2026 GTTCAAGAGAGCTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1266 GTTCAAGAGAGCTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      2086 AGGAGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1326 AGGAGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      2146 GCTCTTCTCCGAGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1386 GCTCTTCTCCGAGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      2206 TGGAGAGAGAGCTTTCAACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1446 TGGAGAGAGAGCTTTCAACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      2266 CTCTTCCGAGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1506 CTCTTCCGAGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      2326 CCGAGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1566 CCGAGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```

RESULT 3

AB134419/c

ID AB134419 standard; DNA; 1608 BP.

AC AB134419;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 2392.

KW Human; immune system disease; cytosine methylation; antiaesthetic;

KW antileukemia; antineoplastic; anti-HIV; anticonvulsant; neurotrophic;

KW antineoplastic; antineoplastic; antidiabetic; ophthalmological;

KW acute myeloid leukemia; Alzheimer's disease; arteriosclerosis; anemia;

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

KW gene; ds.

Claim 1; SEQ ID NO 2392; 32bp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

CC disease: The present sequence is a gene of the invention.

CC Sequence 1608 BP; 366 A; 51 C; 421 G; 770 T; 0 other;

Query Match 29.9%; Score 961; DB 24; Length 1608;

Best Local Similarity 76.6%; Pred. No. 2.5e-251; Indels 2; Gaps 2;

Matches 1202; Conservative 0; Mismatches 365;

767 AGAGATGCGCATATACATCTCTTATCCCACTCTTTTGTCTATATAGATGCA 826

1602 ACAAATAACCTTACCATCTCTTATCCCACTCTTTTATATATATATATCA 1543

827 GAGGTGTGACAGAGCTGCGAGCGCCAGGCTCTCTCTATGCACTACAGTCTGA 886

1542 AATATATACCAAACTACACCGCCAAATCTTCTCTATACATCACTACATCA 1483

887 AGGTGACACAGACAGAGCTATCAACCTACAGCTGTGTGTGTGTGTGTGTGT 1423

1482 AATATATACCAAACTACACCGCCAAATCTTCTCTATACATCACTACATCA 1423

947 ATCAAAACAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1006

1422 ATCAAAACAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1363

1007 CC-AACAGGCTCCAGAGAAATGTCACTGAGAGAGAGAGAGAGAGAGAGAG 1303

1362 CCAACCAATCTCAAAATATCACTTAAATATCACTTAAATATCACTTAAATAT 1243

1066 TAAAGCAAGGT 1185

1302 TAAACCAAAAT 1183

1126 TTGACAGTGTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1123

1242 TTAACCAAAAT 1245

1186 TGTCACTGTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1245

1182 TATCAATCTTAAATATATATATATATATATATATATATATATATATATAT 1305

1246 CCTCTGTAAAT 1063

1122 CCTCTGTAAAT 1365

1306 ACTTGAACAGTGTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1003

1062 ACTTGAACAGTGTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425

1366 GGCAGGAAAT 943

1002 AACCAAAAT 1485

1426 GGGTATGAGTGTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 883

942 AACTATGAGTGTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1545

1486 GGGGCT 823

882 AATCCCT 1605

1546 TCCATTCAGAAAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 763

822 TCCATTCAGAAAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 1665

1606 AATGCGCGCAAGCTTCTCTGAGTACCTGCAATGCAAGGCTCTCTGATACCA 1665

DB 762 AATATCCCGCAATCTTCTCTAATAATCTCAACATATTAACAAATCTCTATACCA 703

1666 GGCAGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCT 1725

702 ACAAATCTCTAGCTCTAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 643

1726 GAATCAACAGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 1785

642 AATCTCAACAAATCTTATATATATATATATATATATATATATATATATAT 583

1786 GTGAGTGTGACAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1845

582 AT-AATATCAACCACTACAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1905

1846 CCTGACT 464

523 CCTTATCT 1965

1906 TGAATCAATGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 404

463 TAAATCAAT 2025

1966 CTGACCAAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 344

403 CTATCAACCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2085

2026 GTTCAATGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 284

343 ATTCATTAATGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2145

2086 AGGGGCGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 224

283 AAAAAAGCAATCTTCAACCAATCTTCAACCAATCTTCAACCAATCTTCAAC 2205

2146 GCT 164

223 ACCCT 2265

2206 TCGAGAGTGTGACAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 104

163 TCGAAATATATCTCAACCAATCTTCAACCAATCTTCAACCAATCTTCAAC 2225

2266 CTCCT 44

103 CTCCT 44

2326 GCGGAGAG 2334

43 GCGGAGAG 35

DB 43 GCGGAGAG 35

RESULT 4

ABL34418 standard; DNA; 1608 BP.

ABL34418;

26-MAR-2002 (first entry)

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

Human immune system associated gene SEQ ID NO: 2391.

30-JUN-2000: 2000DE 102210

1043826.

Olek A, Piepenbrock R

Nucleic acid

Claim 1; SEQ ID NO 2391: 3200

the present invention provides a number of human genes which are modified.

51 C; 423 G; 736 T; 0 other;

conservative
0; Mismatches 371; Indels

19 TATATTTTCTTTTGTCTATTAGATGCTCAGAGTGTGTGCAC 838

839 AAGCTGCGACGCCCAAGGCTTTTCCTCAATGGC
TAAATTATTAAGAATTTAGACTGTGTAT 78

/9 AAGTTGCTACGTTTAGGGTTTTTATGTGAATAACTGTAAGA
| | | | | | | | | | | | | | |
|| || || || || || || || || || || || || ||

898

139 AGATAAGTTTATTTAATTGCTGGTAAGTCAGTGACAGATGC AAAACACA 958

959 GGGTGTGGAAGCCCTCAGGAGGTACCCTAACATATA 198

1018 CAGTAACTTCTGAGCGTTTAAATAGCTTT 353

259 TAGAGAAATGTTAATTCCTCTT
1077

1A1TAATGCTTATTAAATGGGTTCATCTAATTGGGAAGGGAGACACCCCGCCT

1138 GTGCAACGCTTGGAGCGAGAGGTTGTAGAGTGAG 378

379 GTGTAGAGTTGGTGGACGATGTTAAGGAATATTCCTTAAAGAGTCCAGTCTGG 1197

TGATGCAGTTCAGTTCAATTCCAACTGGCGGTAC

1258 TTCATGCTAAACTACCTAAGACGGTATTTCCTTTTGTAAAA 498

499 TTTATGTTAAATTATTTAAGGTATTTATCGTTTTAAACCTT
1317 GGGCCTTGACACAGT

.....GAGCAGAACATTACAAGATTTTTTTTAAAGAAATTGCCACGAATAA
.....

8

610 |||||GAGGACAGGAAGTAATTGTGAATGTTAATATAGCTGGGCTATGCCATT

1438 TGGCTTAAGTTGTTAGCATTGTTCCTTCCCTTGAGAATAATACCT
678

1498 TTCAAGCCTTATGGCGCAACATCTCTT 738

1558 ACCGTCCTCAGCGAATTTTATATTGATTTTATATGGTTAAATGTTATTATTAGA 798

59 ACGTTTGTGAGITTTTATCTGTGACTTATAAGATGCAGTAGTTTACCTTATTTTTCCTT
161

[illegible]

919 GCGTTTGGACCTTTTTCCTCCAGGTGTAATAACAGCAG 1737

|||||TACGAGGGCGCTCTCGAGTCCTCACCCTGAAGTAGCAATCCCA
|||||TTAGAAAGGCAGAGGCCGCTCTCGAGTCCTCACCCTGAAGTAGCAATCCCA

1798 GCCACTGCAGGAACGCCCGGACAGGAATGCCCAATTTCCTC
TAAAG-ATGTTA 1037

1858 CCTCACCCTGACTTCTCCCGCCGCCGCATTTGGATTGTTT 1097

1918 CCCCCCTTTTCTTTTATTCGCCGTAGCTTAAGTTGTGAATTAATGGA 1157

[illegible]

1218 GTTTTCGTGAGTTCGTTTTTGATTTCCCGGGTCTC
2037

1278 TTTTGGGCGGCTGAGCTTGGCTGAGCCCTCAGGGCGGCAC 2097

199CACCAGAACACACCCCTGAGGCCACCCCTGGCTGCCCGAGCGGAGCGTCCCTCTTTT

2158 GCGGTTGGTGACCCGCTCAGTACCGACTTCCCA... 1397

2218 TCAACAAACCCTCGCTT
GAGGTGGCGAAGTTTATTTCGAGGATTC

1AAIAATATGTGGGTAATTGGATTTTATATTTTGGTGAGTTTtttttttttttccccct

1518 TGGAGGTTATTCTGTTTACGTTCCGCGAATCGCT 2337

[illegible]

RESULT 5
AAC98177

[illegible]

QY	Db
2187	121
TTGGGAAAGCTTTTCACTTCGAGCAGTGTCTCAACAACCATGCTCTGCTGCTGAGG	TGGCTGCCCAGCGAGCTGCTCTTCCCGCGGGTGGTGAACCGCTCACTAAGGAG
180	2186

181	TTGGGGAACTCTTTTCACTT	181	TTGCACTCTGGACCCCT	2246
2247	CGTAGGATGCTCAACACCA	2247	CTGCTGGGCACTTGACCCCT	240

241 CCTACTCTTGST 252

RESULT 8
ABV94152
ID ABV94152 standard; cDNA; 2551 bp.
XX
AC ABV94152.

XX
DT
XX
08-JAN-2003 (first entry)

DE	Breast carcinoma related nucleotide sequence SEQ ID NO.143
XX	
MM	

KW breast carcinoma; cancer; tumour; cytostatic; anti-cumour;
 XX gene; ss.

MO200246467-A2.

13-JUN-2002.

P
F
X
07-DEC-2001; 2001WO-IB02811.
R
08-DEC-2000. 2000----

R 07-DEC-2000; 2000US-254090P
X X 2001US-0007926.
A (IPSO-) IPSCC...

Bertucci F, Houligatte R. Birmingham

WPI; 2002-619023/66.

novel polynucleotide library useful in molecular characterization of a carcinoma, comprising a pool of polynucleotide subsequences which are

cells - ... are either underexpressed or overexpressed in tumor

The present invention describes a novel method for the molecular analysis of DNA.

polynucleotide library (I) useful in characterisation of a carcinoma, comprising a pool of polynucleotides or its subsequences which are either overexpressed in tumour cells and

The prognosis or diagnosis of (1) a polynucleotide array, (2) a polynucleotide sequence chosen from the 468 sequences given in ABV94477, and correspond to any of the polynucleotide sequences chosen from the 468 sequences given in ABV94477. Also described is (1) a polynucleotide array, (2) a polynucleotide sequence chosen from the 468 sequences given in ABV94010 (M), and correspond to any of the polynucleotide sequences chosen from the 468 sequences given in ABV94010 (M).

was differentially expressed in tumour, comprising (1) useful for correlated with a cancer, polynucleotide sequences which are a patient, and reacting to the polynucleotide sequences which are

immobilised on a solid support, where the probe comprises any combination of the polynucleotide sequences of (1) or (2) and the probe comprises any combination of polynucleotide sequences of (1) or (2).

The sequences of (I), and detecting the reaction products encoded by cytosolic activities and can be used as anti-tumour agent. It is useful in molecular characterisation of

...tion of a carcinoma. (I) and (II) are agents. (I)

XX
PN
vV
WO200281638-A2.

PD 17-OCT-2002.
XX
XX
PF 08-APR-2002; 2002WO-US10824.
XX
PR 06-APR-2001; 2001US-281731P.
XX
PR 06-APR-2001; 2001US-281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun 2, Jay G;
XX
PI WPI; 2003-058520/05.
XX
PT Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
XX patients -
PS
PS Claim 1; Page 173-174; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
XX prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX sample comprising prostate tissue, which involves determining the number
XX of target genes which are differentially-regulated in the sample, where
XX the number is indicative of the probability that the sample comprises
XX prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX intervention in a subject having a prostate cancer, which involves
XX determining the expression levels in a sample comprising prostate tissue
XX of target genes which are differentially-regulated in prostate cancer.
XX Preferably, the expression levels of at least 10 genes are determined.
XX (I) is also useful for identifying agents that modulate a biological
XX activity of a polypeptide differentially-regulated in prostate cancer.
XX cells, which involves contacting a polypeptide differentially-regulated
XX in prostate cancer cells with a test agent under conditions effective for
XX determining whether the test agent modulates the biological activity of
XX (I) is useful as molecular markers, as drug targets, and for detecting,
XX diagnosing, staging, grading, assessing, monitoring, prognosticating,
XX preventing or treating, determining predisposition to diseases and
XX conditions especially relating to prostate cancer. (I) and its expression
XX products are used in the diagnostic test to assay for presence of cancer
XX e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
XX of cancer, its stage of development, the nature of genetic defect, etc.
XX The polypeptide encoded by (I) can be used as target for therapy or drug
XX discovery. (I) can also be used as target for therapy or drug
XX for searching specific binding partners of the polypeptide and thus
XX useful in therapeutic applications to treat prostate cancer. (I) is
XX identification of specific genes, and groups of genes, expressed in
XX pathways physiologically relevant to prostate cancer permits the
XX definition of functional and disease pathways and the delineation of
XX targets in these pathways which are useful in diagnostic, therapeutic,
XX and clinical applications. This sequence encodes a protein differentially
XX regulated in prostate cancer.
SQ
Sequence 2551 BP; 826 A; 493 C; 508 G; 724 T; 0 other;
Query Match 7.8%; Score 252; DB 25; Length 2551;
Best Local Similarity 100.0%; Pred. No. 1.2e-57;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2007 GCAAGAGTGCACACAGGTTTCAAGAGCGCTTGGGAGTGGAGGAGCGGTTTACGA 2066
DB 1 GCAAGAGTGCACACAGGTTTCAAGAGCGCTTGGGAGTGGAGGAGCGGTTTACGA 2066
QY 2067 GTGACTGTGCTGAGAGCTTCAAGGAGGAGGAGTGGAGGAGCGGTTTACGA 60
DB 61 GTGACTGTGCTGAGAGCTTCAAGGAGGAGGAGTGGAGGAGCGGTTTACGA 60
QY 2127 TGGCTGCCAGGCGGAGCTGCTTCTTCCGCGGAGTTTGGAGCCGCTCAGTACGAG 2186
DB 121 TGGCTGCCAGGCGGAGCTGCTTCTTCCGCGGAGTTTGGAGCCGCTCAGTACGAG 180

QY 2187 TTGGGAGAGCTTTTCACTTGGAGATTGCTCAACACCATGCGGAGCTTGGACCT 2246
DB 181 TTGGGAGAGCTTTTCACTTGGAGATTGCTCAACACCATGCGGAGCTTGGACCT 2246
QY 2247 CCTACCTTCGCT 2258
DB 241 CCTACCTTCGCT 252
RESULT 11
ABZ83782
ID ABZ83782 standard; cDNA; 1024 BP.
XX
AC ABZ83782;
XX
DT 14-MAY-2003 (first entry)
XX
DE Toxicologically relevant human nucleotide sequence #941.
XX
KW Toxicologically relevant gene; toxicological response; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003016500-A2.
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-US26514.
XX
PR 16-AUG-2001; 2001US-313080P.
XX
PA (PHAS-) PHAS-1 MOLECULAR TOXICOLOGY INC.
XX
PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeisler K;
PI Alen P;
XX
DR WPI; 2003-268322/26.
XX
PT Determining a toxicological response to an agent, useful for screening
PT of drugs, comprises comparing the expression profile of one or more
PT human toxic response genes to a reference gene expression profile
XX indicative of toxicity -
XX
PS Claim 1; Page 294; 455pp; English.
XX
XX The present invention describes a method (M1) for determining a
XX toxicological response to an agent which comprises comparing a
XX reference gene expression profile of one or more human toxic response genes to a
XX determining the presence of a toxic response to the agent. Also
XX from the genes corresponding to the partial sequences given in ABZ82842
XX to ABZ4764, or their fragments of at least 20 nucleotides, or
XX homologues; and (2) determining if a gene putatively identified by
XX determining the expression profile of the gene after exposure of cells
XX or a human subject to a known toxic pharmaceutical or industrial agent,
XX comprising: (a) exposing cells to an agent or isolating cells from a
XX human subject who was exposed to an agent; (b) obtaining the test gene
XX expression profile for a putatively identified toxic response gene after
XX exposure to a known toxic pharmaceutical or industrial agent; and
XX (c) comparing the test profile to the expression profile of a gene with
XX a similar function or comparing the test profile to the expression
XX profile of that gene after exposure to other known toxic compounds. The
XX methods are useful for predicting and determining toxicological responses
XX on a cellular, organ or system level. The arrays comprising the human
XX genes are useful for toxicological screening of drugs, pharmaceutical
XX compounds and chemicals.
SQ
Sequence 1024 BP; 308 A; 226 C; 256 G; 234 T; 0 other;
Query Match 7.0%; Score 226; DB 25; Length 1024;
Best Local Similarity 100.0%; Pred. No. 8.4e-51;

Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGGTTTACAGTGAAGCTTGAGAGCTTCAGGGGCG 2092
1 GACGCTTCGGGGAGTGAAGGAGCGGTTTACAGTGAAGCTTGAGAGCTTCAGGGGCG 60

DB 2093 GGCACCTGGACGGAACACACCTCTGAGGCGACCCCTGGGCTGCCAGGCGAGCTGCTCTT 2152
61 GGCACCTGGACGGAACACACCTCTGAGGCGACCCCTGGGCTGCCAGGCGAGCTGCTCTT 120

QY 2153 CTCCTGGCGGGTGTGGTGAAGCCCGCTCAGTACGAGTGGGAGAGCTTTCACCTTCGAGG 2212
121 CTCCTGGCGGGTGTGGTGAAGCCCGCTCAGTACGAGTGGGAGAGCTTTCACCTTCGAGG 180

DB 2213 ATTGCTCAACACACATGCTGGGCACTTGAGACCTCTCACTCTGGT 2258
181 ATTGCTCAACACACATGCTGGGCACTTGAGACCTCTCACTCTGGT 226

RESULT 12

AAQ93879 ID AAQ93879 standard; cDNA; 2471 BP.

XX AC AAQ93879;

XX 25-MAR-2003 (updated)

DT 06-NOV-1995 (first entry)

XX FAS-DELTA-TM CDNA.

XX FAS-DELTA-TM; transmembrane deletion; apoptosis; antibody;
KW adoptive immunotherapy; transgenic animal; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 195..1139

XX sig_peptide /tag= a

XX mat_peptide /tag= b

XX /tag= c

XX MO9513701-A1.

XX 26-MAY-1995.

XX 15-NOV-1994; 94WO-US13173.

XX 15-NOV-1993; 93US-0152443.

XX (LXRB-) LXR BIOTECHNOLOGY INC.

XX Barr PJ, Kiefer MC, Shapiro JP;

XX WPI; 1995-200120/26.

XX P-PSDB; AAR76238.

XX New nucleic acid encoding Fas protein without its trans-membrane region

XX and related vectors, transformed cells, transgenic animals, protein and

XX antibodies, useful for control of Fas mediated apoptosis

XX Claim 3; Fig.3-1 to 3-4; 38pp; English.

XX mRNA was obt'd. from human lymphocytes and PCR was used to make

XX cDNA specific for Fas-delta-TM (i.e. Fas lacking the transmembrane

XX region) mRNA. The PCR product was ligated into pBluescript and the

XX recombinant plasmid was used to transfect E. coli DH5-alpha cells. The

XX insert sequence of pBluescript-Fas-delta-TM is given in AAQ93879.

XX (updated on 25-MAR-2003 to correct PN field.)

XX CC

XX Sequence 2471 BP; 807 A; 474 C; 489 G; 701 T; 0 other;

Query Match 7.0%; Score 226; DB 16; Length 2471;
Best Local Similarity 100.0%; Pred. No. 1.5e-50;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGGTTTACAGTGAAGCTTGAGAGCTTCAGGGGCG 2092
1 GACGCTTCGGGGAGTGAAGGAGCGGTTTACAGTGAAGCTTGAGAGCTTCAGGGGCG 60

DB 2093 GGCACCTGGACGGAACACACCTCTGAGGCGACCCCTGGGCTGCCAGGCGAGCTGCTCTT 2152
61 GGCACCTGGACGGAACACACCTCTGAGGCGACCCCTGGGCTGCCAGGCGAGCTGCTCTT 120

QY 2153 CTCCTGGCGGGTGTGGTGAAGCCCGCTCAGTACGAGTGGGAGAGCTTTCACCTTCGAGG 2212
121 CTCCTGGCGGGTGTGGTGAAGCCCGCTCAGTACGAGTGGGAGAGCTTTCACCTTCGAGG 180

DB 2213 ATTGCTCAACACACATGCTGGGCACTTGAGACCTCTCACTCTGGT 2258
181 ATTGCTCAACACACATGCTGGGCACTTGAGACCTCTCACTCTGGT 226

RESULT 13

AAK24878 ID AAK24878 standard; DNA; 2471 BP.

XX AAK24878;

XX 21-JUN-1999 (first entry)

DT 21-JUN-1999 (first entry)

XX Soluble Fas receptor DNA.

XX Fas receptor; Fas ligand; FasL; proinflammatory; immunosuppressive;

KW graft versus host disease; autoimmune disease; psoriasis;

KW rheumatoid arthritis; systemic lupus erythematosus; gene therapy;

XX ss.

XX Mammalia.

XX Key Location/Qualifiers

XX CDS 195..1139

XX sig_peptide /tag= a

XX /transl_except= (pos:519..521, aa:Gly)

XX mat_peptide /tag= b

XX polyA_signal /tag= c

XX /tag= d

XX /tag= e

XX MO9903999-A1.

XX 28-JUN-1999.

XX 16-JUL-1998; 98WO-US14771.

XX 17-JUL-1997; 97US-0052829.

XX (UNMI) UNIV MICHIGAN.

XX Chen J, Nabel GJ;

XX WPI; 1999-132243/11.

XX P-PSDB; AAW98070.

XX Inhibition of proinflammatory responses - using an agent which

XX modulates FasL stimulation, used for treating graft versus host

XX disease or autoimmune disease

XX Disclousure; Fig 4B; 71pp; English.

XX This present sequence is a DNA clone encoding soluble Fas receptor

XX CC

(see AAW98070). The invention provides a method for inhibiting a proinflammatory response in a cell mixture by administering an immunosuppressive agent which inhibits the proinflammatory activity of Fas ligand (FasL). In some embodiments, FasL is coadministered with the immunosuppressive agent, and the cell mixture comprises neutrophil cells. The method can be practiced in vitro, ex vivo or in vivo. Suitable immunosuppressive agents include antisense molecules that inhibit endogenous FasL expression, soluble Fas receptors, ribozymes that inhibit the endogenous expression of FasL, drugs that inhibit FasL signaling, agents that induce of endogenous expression of transforming growth factor (TGF)-beta, and polynucleotides coding for an immunosuppressive agent such as TGF-beta. The method can be used for treating diseases associated with an undesired FasL-mediated proinflammatory response, such as graft versus host disease, or an autoimmune disease such as systemic lupus erythematosus, rheumatoid arthritis and psoriasis. The invention also provides a method for identifying agents which modulate FasL stimulation of a proinflammatory response.

Sequence 2471 BP; 807 A; 474 C; 489 G; 701 T; 0 other;

Query Match
Best Local Similarity 7.0%; Score 226; DB 20; Length 2471;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCTCAGGGCG 2092
DB 1 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCTCAGGGCG 60
QY 2093 GGCACCTGGCAGGACACACACCTGAGGCGACCTGGCTGCCAGGGAGTGCCTCTT 2152
DB 61 GGCACCTGGCAGGACACACACCTGAGGCGACCTGGCTGCCAGGGAGTGCCTCTT 120
QY 2153 CTCGGCGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGAGCTCTTCACTTCGGAGG 2212
DB 121 CTCGGCGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGAGTCTTTACTTCGGAGG 180
QY 2213 ATTGCTCAACCAACCATCTGGGACATCTGAGACCTCTCACTCTGCT 2258
DB 181 ATTGCTCAACCAACCATCTGGGACATCTGAGACCTCTCACTCTGCT 226

RESULT 14
AAQ29959 standard; cDNA to mRNA; 2534 BP.

AC AAQ29959;
XX 25-MAR-2003 (updated)
DT 12-MAR-1993 (first entry)
XX
DE Human cell surface antigen.
XX
KW Fas antigen; apoptosis; pF58; NGFR/TNFR family; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT sig_peptide /tag= a
FT mat_peptide /tag= b
FT /product= Fas_antigen
FT conflict /tag= c
FT /note= "this residue is not present in pf3"
FT polyA_signal /tag= d
FT polyA_signal /tag= e
FT polyA_signal /tag= e
FT 2518..2523
XX /tag= f

PN BP510691-A1.
XX 28-OCT-1992.
XX
PF 24-APR-1992; 92EP-0107060.
XX
PR 26-APR-1991; 91JP-0125234.
XX
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX
PI Itoh N, Nagata S, Yonehara S;
XX
DR WPI, 1992-358914/44.
XX
P-PSDB; AAR28084.
XX
PT DNA encoding human cell surface antigen - used to clarify
XX apoptosis mechanism of various types of cell, and to prepare
XX monoclonal antibodies that react with tumour cells expressing Fas
PS Claim 3; Fig 1 and 2; 27pp; English.

A cDNA library was prepared from polyA+ RNA from the human lymphoma cell line KT-3. The cDNA was ligated to BstXI-cut vector pCEV4 via BstXI linkers. The KT3 cDNA library was used to transfect monkey COS-7 cells which were then suspended in buffer containing murine anti-Fas antibodies. The Fas-expressing cells adhered to the plates. Extrachromosomal DNA was prepared from adhered cells and used to transform E.coli WMI00 cells. A 520bp XhoI-BamHI fragment from a positive clone (pF3) was used to screen the KT-3 cDNA library. The longest cDNA clone was designated pF58 and contains an ORF corresp. to a 335 amino acid pre-protein and a 319 amino acid mature protein (1.e. human Fas antigen).

Sequence 2534 BP; 817 A; 487 C; 503 G; 727 T; 0 other;

Query Match
Best Local Similarity 7.0%; Score 226; DB 13; Length 2534;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCTCAGGGCG 2092
DB 1 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCTCAGGGCG 60
QY 2093 GGCACCTGGCAGGACACACACCTGAGGCGACCTGGCTGCCAGGGAGTGCCTCTT 2152
DB 61 GGCACCTGGCAGGACACACACCTGAGGCGACCTGGCTGCCAGGGAGTGCCTCTT 120
QY 2153 CTCGGCGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGAGCTCTTCACTTCGGAGG 2212
DB 121 CTCGGCGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGAGTCTTTCACTTCGGAGG 180
QY 2213 ATTGCTCAACCAACCATCTGGGACATCTGAGACCTCTCACTCTGCT 2258
DB 181 ATTGCTCAACCAACCATCTGGGACATCTGAGACCTCTCACTCTGCT 226

RESULT 15
AAQ95297 standard; cDNA; 2534 BP.

AC AAQ95297;
XX 19-FEB-1996 (first entry)
DT 19-FEB-1996 (first entry)
XX
DE Plasmid pF58 contg. human Fas cDNA.
XX
KW Plasmid pF58; human Fas cDNA; soluble membrane protein;
XX antibody production; diseases; treatment; prevention; ds.
OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	195..1202
FT sig_peptide	/*tag= a
FT mat_peptide	195..242
FT	/*tag= b
FT	243..1199
FT	/*tag= c

JP07115988-A.

XX 09-MAY-1995.

XX 26-OCT-1993; 93JP-0267644.

XX 26-OCT-1993; 93JP-0267644.

XX (NISB) JAPAN TOBACCO INC.

XX WPI; 1995-202847/27.

XX P-PSDB; AAR78606.

XX Preparation of soluble membrane proteins - for their use in antibody

XX production for the treatment and prevention of related diseases

XX Example 1; Pages 15-17; 51pp; Japanese.

XX AAQ95297 is the plasmid pF58 which contains the human Fas CDNA. The

XX plasmid was used in the construction of an expression vector for

XX the prodn. of recombinant soluble membrane proteins. The proteins

XX can be used in antibody prodn. for the treatment and prevention of

XX related diseases.

XX Sequence 2534 BP; 817 A; 487 C; 503 G; 727 T; 0 other;

XX Query Match 7.0%; Score 226; DB 16; Length 2534;

XX Best Local Similarity 100.0%; Pred. No. 1.5e-50;

XX Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

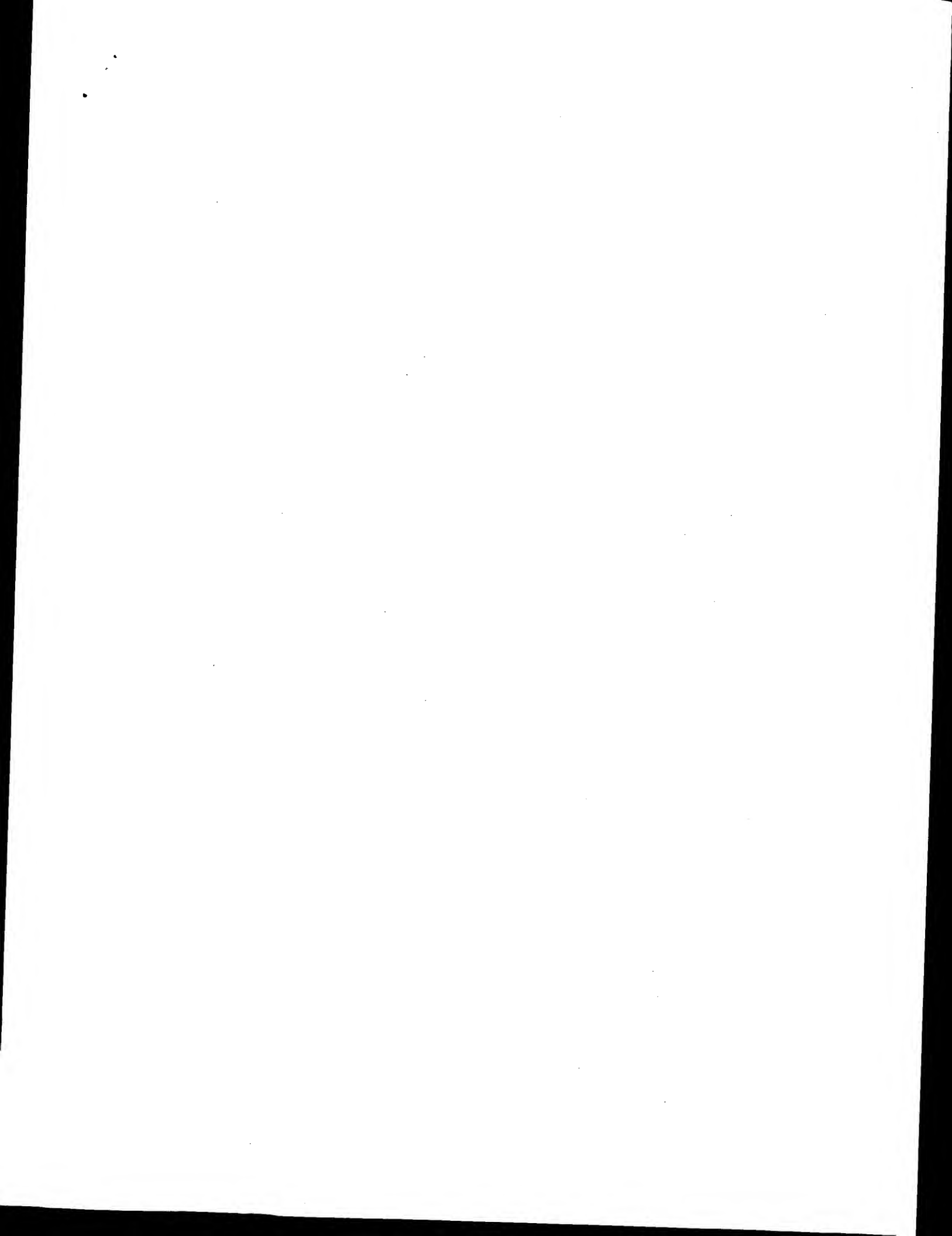
XX

XX

XX

XX

Search completed: August 1, 2003, 13:55:11
 Job time : 1079.81 secs



Dp	2341	GGATTCCGGCGGACGGCGGCA	CGCGGGCACCTGGGA	CGCGGGCTGCTGCGGAGGGG	2400	
Qy	2341	GGATTCCGGCGGACGGCGGCA	CGCGGGCACCTGGGA	CGCGGGCTGCTGCGGAGGGG	2400	
Dp	2401	TTGGAGACTGGCTCCCGGGGGCTGT	TAGAACCTTCCCTCAGGCCGGGGTGT	CGAAGCA	2460	
Qy	2401	TTGGAGACTGGCTCCCGGGGGCTGT	TAGAACCTTCCCTCAGGCCGGGGTGT	CGAAGCA	2460	
Dp	2401	TTGGAGACTGGCTCCCGGGGGCTGT	TAGAACCTTCCCTCAGGCCGGGGTGT	CGAAGCA	2460	
Qy	2461	TGGAGGACTTGTCTTTCTTGTGG	CGCTTGAATGCGAAGCTGATCCG	CTGGGCAAGCGGGG	2520	
Dp	2461	TGGAGGACTTGTCTTTCTTGTGG	CGCTTGAATGCGAAGCTGATCCG	CTGGGCAAGCGGGG	2520	
Qy	2521	CGAGTCGGCGGCTCTCGGAGACA	CTGCGCTCCAGCTTGAAGTGGGCG	CTGGGGCGGCA	2580	
Dp	2521	CGAGTCGGCGGCTCTCGGAGACA	CTGCGCTCCAGCTTGAAGTGGGCG	CTGGGGCGGCA	2580	
Qy	2521	CGAGTCGGCGGCTCTCGGAGACA	CTGCGCTCCAGCTTGAAGTGGGCG	CTGGGGCGGCA	2580	
Dp	2581	CAGGAATTGAACCGGAAGCTG	GGAGAGCTTTAGGTCGTGAGAGGGG	CAACCCGGTTGA	2640	
Qy	2581	CAGGAATTGAACCGGAAGCTG	GGAGAGCTTTAGGTCGTGAGAGGGG	CAACCCGGTTGA	2640	
Dp	2581	CAGGAATTGAACCGGAAGCTG	GGAGAGCTTTAGGTCGTGAGAGGGG	CAACCCGGTTGA	2640	
Qy	2641	GAGAGAGCGGAACTCCTTGA	CAAGCCCTGACAAAGCCAAAGCT	CCGCTCCGGCGC	2700	
Dp	2641	GAGAGAGCGGAACTCCTTGA	CAAGCCCTGACAAAGCCAAAGCT	CCGCTCCGGCGC	2700	
Qy	2701	GGGTGGGTGAGTGC	CGCGCGCGCGCGGGGGCGGGAGAGGCTT	CAGCTTCAAGCA	2760	
Dp	2701	GGGTGGGTGAGTGC	CGCGCGCGCGCGCGGGGGCGGGAGAGGCTT	CAGCTTCAAGCA	2760	
Qy	2761	CATATTGCTCATTTTCTGG	CACTTCTGAGAGCTTGAAGTAAGT	CAGACCGAAGCA	2820	
Dp	2761	CATATTGCTCATTTTCTGG	CACTTCTGAGAGCTTGAAGTAAGT	CAGACCGAAGCA	2820	
Qy	2821	GTTAAGCCGGAGGCTCGAGAG	ACGGCACTTTCTTTCGAAAAGTTAT	TGGGGG	2880	
Dp	2821	GTTAAGCCGGAGGCTCGAGAG	ACGGCACTTTCTTTCGAAAAGTTAT	TGGGGG	2880	
Qy	2881	CTGAATGAGCTTCTGGAG	AGCTTGTTAACGTTTTTATGTCAC	CAGAAAAGAACTG	2940	
Dp	2881	CTGAATGAGCTTCTGGAG	AGCTTGTTAACGTTTTTATGTCAC	CAGAAAAGAACTG	2940	
Qy	2941	CTTGTCTCCCTTCCGGAA	ATTCTCTTGAAGCTGTAAGTGC	GTGCGTGAATGCTG	3000	
Dp	2941	CTTGTCTCCCTTCCGGAA	ATTCTCTTGAAGCTGTAAGTGC	GTGCGTGAATGCTG	3000	
Qy	2941	CTTGTCTCCCTTCCGGAA	ATTCTCTTGAAGCTGTAAGTGC	GTGCGTGAATGCTG	3000	
Dp	3001	ATTGTGTTTGTTTTCTG	CCCTTCTCTTTCGCTTTCGCTT	CTTGAAGCTGCG	3060	
Qy	3001	ATTGTGTTTGTTTTCTG	CCCTTCTCTTTCGCTTTCGCTT	CTTGAAGCTGCG	3060	
Dp	3061	CATGTGATTTCTGCTTGT	CTCTGTGGAGTTGTGTATCGT	TCCACCGCAGCA	3120	
Qy	3061	CATGTGATTTCTGCTTGT	CTCTGTGGAGTTGTGTATCGT	TCCACCGCAGCA	3120	
Dp	3121	ACCGGCGGCTATTTATG	GCACAAAGAACTTGAGCA	AGCTTTTGA	AAAGTCCCTGCTC	3180
Qy	3121	ACCGGCGGCTATTTATG	GCACAAAGAACTTGAGCA	AGCTTTTGA	AAAGTCCCTGCTC	3180
Dp	3121	ACCGGCGGCTATTTATG	GCACAAAGAACTTGAGCA	AGCTTTTGA	AAAGTCCCTGCTC	3180
Qy	3181	AGAAATGCAAGCTTGCA	ATGAGTGCATATCAAG	3212		
Dp	3181	AGAAATGCAAGCTTGCA	ATGAGTGCATATCAAG	3212		
Qy	3181	AGAAATGCAAGCTTGCA	ATGAGTGCATATCAAG	3212		

RESULT 2
 US-09-834-291-4
 Sequence 4, Application US/09834291
 Patent No. US20020042064A1
 GENERAL INFORMATION:
 APPLICANT: Krammer, Peter
 APPLICANT: Muller-Schilling, Martina
 APPLICANT: Oren, Moshe
 TITLE OF INVENTION: p53 Binding Areas
 FILE REFERENCE: 4121-122
 CURRENT APPLICATION NUMBER: US/09/834,299
 CURRENT FILING DATE: 2001-08-21
 PRIORITY APPLICATION NUMBER: PCT/DE99/03543

; PRIOR FILING DATE: 1999-10-18
 ; PRIOR APPLICATION NUMBER: DE 198 47 779.1
 ; PRIOR FILING DATE: 1998-10-16
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2827
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; US-09-834-291.4

[illegible]

841 GCGTGCACGCCAGGGCTTCTCATGCGCACTAACGCTACTGAAAGGTGAAACAGAG 900
QY 901 ACAAGCCCTATCAAACTTACAGAGCTGCTGTAAGTGCATGACATGCAAAAACAGAG 960
Db 901 ACAAGCCCTATCAAACTTACAGAGCTGCTGTAAGTGCATGACATGCAAAAACAGAG 960
QY 961 GTATGAAAGCCCTCAGAGAGGTTAACTTAACTTGAATTTAGAGGCCC-AACAGCTCCA 960
Db 961 GTATGAAAGCCCTCAGAGAGGTTAACTTAACTTGAATTTAGAGGCCC-AACAGCTCCA 960
QY 1020 GAAGAAATATGCACTGAGAGAGAGCTGAAAGATGAAAGTGGCTTAAAGCAAGGTTA 1079
Db 1021 GAAGAAATATGCACTGAGAGAGAGCTGAAAGATGAAAGTGGCTTAAAGCAAGGTTA 1079
QY 1081 TTAATGTGTTAATTAATGAGTTGAATCTAATTTGGAAGAGAGAGAGGTTA 1080
Db 1081 TTAATGTGTTAATTAATGAGTTGAATCTAATTTGGAAGAGAGAGAGGTTA 1080
QY 1140 GCAAGCTTGTGAGAGAGAGCTGAAAGATGAAAGTGGCTTAAAGCAAGGTTA 1140
Db 1141 GCAAGCTTGTGAGAGAGAGCTGAAAGATGAAAGTGGCTTAAAGCAAGGTTA 1140
QY 1200 CTGATCCAAATTCAGGTTCAATATGATGATATCAATATCAATCTTGTAAATTT 1259
Db 1201 CTGATCCAAATTCAGGTTCAATATGATGATATCAATATCAATCTTGTAAATTT 1259
QY 1260 CATGCTAAATTAACCTTAAGAGTATCTACCGTCCAAAGCAATGTAAGTGAACAGTGT 1260
Db 1261 CATGCTAAATTAACCTTAAGAGTATCTACCGTCCAAAGCAATGTAAGTGAACAGTGT 1260
QY 1320 TCACCAAGAGAGAGAGATTAAGAAATTTTAAAGAAATTTGAGCAAGAAATAT 1379
Db 1321 TCACCAAGAGAGAGAGATTAAGAAATTTTAAAGAAATTTGAGCAAGAAATAT 1379
QY 1380 GAGTAAAGAGAGAGAGAGATTAAGTAAATTTTAAAGAAATTTGAGCAAGAAATAT 1380
Db 1381 GAGTAAAGAGAGAGAGAGATTAAGTAAATTTTAAAGAAATTTGAGCAAGAAATAT 1380
QY 1440 GCTTAAGTTGTTAGTTGTTTCTCTTGAAGAAATTAAGTAAATTTGAGCAAGAAATAT 1440
Db 1441 GCTTAAGTTGTTAGTTGTTTCTCTTGAAGAAATTAAGTAAATTTGAGCAAGAAATAT 1440
QY 1500 CAGAGCCCTTATGAGAGAGAGATTAAGTAAATTTTAAAGAAATTTGAGCAAGAAATAT 1500
Db 1501 CAGAGCCCTTATGAGAGAGAGATTAAGTAAATTTTAAAGAAATTTGAGCAAGAAATAT 1500
QY 1560 GTCTGTAGAGCTCTCATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1561 GTCTGTAGAGCTCTCATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1620 TCTTTCTGAG 1620
Db 1621 TCTTTCTGAG 1620
QY 1680 GCTGTAGAGCTCTCATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1681 GCTGTAGAGCTCTCATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1740 GCTTTTGAAG 1740
Db 1741 GCTTTTGAAG 1740
QY 1800 CACTGAG 1800
Db 1801 CACTGAG 1800
QY 1860 TCACCTGAG 1860
Db 1861 TCACCTGAG 1860
QY 1920 CTTCCCAAG 1920
Db 1921 CTTCCCAAG 1920
QY 1979 CTTCCCAAG 1979
Db 1980 CTTCCCAAG 1980

1980 TTTTGTGAG 2039
QY 1981 TTTTGTGAG 2039
Db 1981 TTTTGTGAG 2039
QY 2040 CTGGAG 2040
Db 2041 CTGGAG 2040
QY 2100 GCAAG 2100
Db 2101 GCAAG 2100
QY 2160 GGTGTGAG 2160
Db 2161 GGTGTGAG 2160
QY 2220 AACAG 2220
Db 2220 AACAG 2220
QY 2188 GAGGCTTACCCGCTCTAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2187
Db 2188 GAGGCTTACCCGCTCTAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2187
QY 2188 GAGGCTTACCCGCTCTAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2187
Db 2188 GAGGCTTACCCGCTCTAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2187
QY 2340 GGAATGAG 2340
Db 2340 GGAATGAG 2340
QY 2400 GTTGAAG 2400
Db 2400 GTTGAAG 2400
QY 2460 ATGAG 2460
Db 2460 ATGAG 2460
QY 2520 GCAAGCTCCGAG 2520
Db 2520 GCAAGCTCCGAG 2520
QY 2580 ACAG 2580
Db 2580 ACAG 2580
QY 2640 AG 2640
Db 2640 AG 2640
QY 2700 CCGGTGAG 2700
Db 2700 CCGGTGAG 2700
QY 2760 ACATATGAG 2760
Db 2760 ACATATGAG 2760
QY 2820 GGTAAAG 2820
Db 2820 GGTAAAG 2820
QY 2880 GGTAAAG 2880
Db 2880 GGTAAAG 2880
QY 2940 GGTAAAG 2940
Db 2940 GGTAAAG 2940
QY 3000 CATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3000
Db 3000 CATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3000
QY 3060 CATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3060
Db 3060 CATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3060

[illegible]

QY	867	TGGCACTAACAGTCTACAGAAAGGTGGACAGAGACAGACCTTACCAACCTACCAAGCT	926
Db	421	TGGCACTAACAGTCTACAGAAAGGTGGACAGAGACAGACCTTACCAACCTACCAAGCT	480
QY	927	GGGTGTAAGTGCAGTGCAGATGCMAAACACAGGGTATGAAAAGCCCTCAGAGGGTAA	985
Db	481	GGGTGTAAGTGCAGTGCAGATGCMAAACACAGGGTATGAAAAGCCCTCAGAGGGTAA	540
QY	987	CCPAACCTAATTTGAGGGGCC-ACAGGCTCCAGAGAAATGTCAATCAGAGAGAC	1045
Db	541	CCPAACCTAATTTGAGGGGCC-ACAGGCTCCAGAGAAATGTCAATCAGAGAGAGC	600
QY	1046	CTGAAGATGAAACAGTGGGCTAGCAAAAGGTTATTAATGTGTTATTAATGGGTGATC	1105
Db	601	CTGAAGATGAAACAGTGGGCTAGCAAAAGGTTATTAATGTGTTATTAATGGGTGATC	660
QY	1106	TAATTGGAAAGGAGAGAGGTTGCAGAGTGCAGAGCTTGTGTGACGATGCCAAG	1165
Db	661	TAATTGGAAAGGAGAGAGGTTGCAGAGTGCAGAGCTTGTGTGACGATGCCAAG	720
QY	1166	GAATACGAAACCTTAGTGTGTGCAAGCTTGAACTGCATCCAAATTCAGTTCCAGTAT	1225
Db	721	GAATACGAAACCTTAGTGTGTGCAAGCTTGAACTGCATCCAAATTCAGTTCCAGTAT	780
QY	1226	GATGCTATATCCAAACATACCTTGTAATTAATTCATGCTAACTACCPAAGGCTATCT	1285
Db	781	GATGCTATATCCAAACATACCTTGTAATTAATTCATGCTAACTACCPAAGGCTATCT	840
QY	1286	ACCGTCCAAAGCAATAGTGACTTTGACAGGTTCACACAGACAGAAATTAACAG	1345
Db	841	ACCGTCCAAAGCAATAGTGACTTTGACAGGTTCACACAGACAGAAATTAACAG	900
QY	1346	ATTTTTTTTAAAGAAATTTGACAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	1405
Db	901	ATTTTTTTTAAAGAAATTTGACAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	960
QY	1406	TGAATGTTAATATATAGCTGGGGCTATGCAATTTGGCTTAAGTTGTTGTTCT	1465
Db	961	TGAATGTTAATATATAGCTGGGGCTATGCAATTTGGCTTAAGTTGTTGTTCT	1020
QY	1466	CTTGAGAAATTAATACTAAAGGGCCCTCTTTCAAGCTTATGAGCGCAATCTGTA	1525
Db	1021	CTTGAGAAATTAATACTAAAGGGCCCTCTTTCAAGCTTATGAGCGCAATCTGTA	1080
QY	1526	CTTTTCAATATGTTAATGCTCAATTCAGAAAGCTGTGAGGCTCTCATGTTCAGCC	1585
Db	1081	CTTTTCAATATGTTAATGCTCAATTCAGAAAGCTGTGAGGCTCTCATGTTCAGCC	1140
QY	1586	ACAACATGACAGCCCAATGCAATGCCCCGCAAGCTTTCTGATGATCTCCAGCAAT	1645
Db	1141	ACAACATGACAGCCCAATGCAATGCCCCGCAAGCTTTCTGATGATCTCCAGCAAT	1200
QY	1646	AGCAGAGGCTCTGTAACCCAGAGAGACCTGTGGCTGTGAAGTCCATCTCTTCAGGA	1705
Db	1201	AGCAGAGGCTCTGTAACCCAGAGAGACCTGTGGCTGTGAAGTCCATCTCTTCAGGA	1260
QY	1706	CCCTCCCACTTCCAGAGTGAACACAGCAGAAAGCTTTGAAAGGAGAGAGCCGAC	1765
Db	1261	CCCTCCCACTTCCAGAGTGAACACAGCAGAAAGCTTTGAAAGGAGAGAGCCGAC	1320
QY	1766	TCTGAGAGTCTCACTGAAGTGAACATGCAAGCACTGCAAGAAAGCCCGGAGACAGA	1825
Db	1321	TCTGAGAGTCTCACTGAAGTGAACATGCAAGCACTGCAAGAAAGCCCGGAGACAGA	1380
QY	1826	ATGCCCATTTGTGCAAGAAACCTGACCTCTTCTCAACCTGCTCCCTCCCTAC	1885
Db	1381	ATGCCCATTTGTGCAAGAAACCTGACCTCTTCTCAACCTGCTCCCTCCCTAC	1440
QY	1886	CCGCGCGAGCGCAAGTTGCTGAATCAATGAACCTCCCAACCGGCGGTTCGCCAGC	1945
Db	1441	CCGCGCGAGCGCAAGTTGCTGAATCAATGAACCTCCCAACCGGCGGTTCGCCAGC	1500
QY	1946	GAGGCTTCCCTCCATCTCTCTGACACCGGAGGCTTTGTGAGCTGCTGATCTCG	2005

conservative 0; Mismatches 24; Indels 394; Cons

1 AGCTTTTGGTACA TTTTTTTT TTTTATGAGTAAATCACTCATCTCACTG 507

61 GGCCTATAATCCTTAACTTTGAGGACGGAGAACCCACATATGTGAGTTGCTGGCTTATAAT 567

131
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

|||||||CTTCCCTTACCTCTCCTTTCTCTCCCTGACACACCCTTTTCCTGCCm

000 CTTTACATTTTATTTAATGAACTTTTCATTTTGAATACTTTTCTTTT

748 AAATTTCAGAGATTAATACAGAGAAATGCCCATATACCAATGCGAATTTT
|||

808 TGTGCTATTAGATGCTCAGAGTGTCCTCTTCTTTT 360

868 GGCACTACAGTCTATTCAAGGCTTCTCAT 420

928 GTGGTAACTCCACGCGAAGCCTATCAACACCCTACAAGACTG 480

988 CTTAAGCAAGGCTGATGGAAGCCCTCAGGAGGCTAAC 540

.....CCAGGAIITGAGGGCCCCAAACAGGCTCCAGAGAATAATGTCAACTGAGAGCATTCCC

1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818

661 AATTGGAGCGAGAGGTTGCAGAGTACCTCCTCCATGAGG 1166

721 AATACTGAACCTTAA GTGTGTCCTCCTCCATCAAGTT CAGGTTCAGTAATG 1226

781 ATGTCATTATCCAAACCTAGAGCTATCTA 1286

841
CGGTTCCAGTCAATTAACAAGA 1346

|||||TGGCAGGAATTAATGACTAACGAGGCACAGGAAGTAATTGT 1405
|||||

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

178 / 1 GAGAAATAAAACTAAGGGCCCTCCCTTTTCAGAGCCTTATGCCCGG

Db 1021 TTGAGAAATAAAACTAAGGGGCTCCCTTTTCAAGAGCCCTATGGCGCAACTCTGAC 1080
 Qy 1527 TTTTTCATANGTTAACTGTCATTTCCAGAAAGCTCTGTAGAGCTCTCATGTTGAGCAAC 1586
 Db 1081 TTTTTCATANGTTAACTGTCATTTCCAGAAAGCTCTGTAGAGCTCTCATGTTGAGCAAC 1140
 Qy 1587 CAACATGACAGCCGACATCAATGATGCGGCAAGCTCTTCTCTGATGATCTCCAGCAATTA 1646
 Db 1141 CAACATGACAGCCGACATCAATGATGCGGCAAGCTCTTCTCTGATGATCTCCAGCAATTA 1200
 Qy 1647 GCGAAGGCTCTGTATCCAGAGGAGCACTCTGCGCTCTGAGACTCATTTCTCTCAAGAC 1706
 Db 1201 GCGAAGGCTCTGTATCCAGAGGAGCACTCTGCGCTCTGAGACTCATTTCTCTCAAGAC 1260
 Qy 1707 CTCCCAACTTCCAGGTTGAACTACAGCAAGGCTTTAGAAAGGCAAGAGCGGCT 1766
 Db 1261 CTCCCAACTTCCAGGTTGAACTACAGCAAGGCTTTAGAAAGGCAAGAGCGGCT 1320
 Qy 1767 CTGAGGCTCTCACTGAGAGTGAAGTGCAGCACTGAGAGAGAGCGCCGCGGAGAGAA 1826
 Db 1321 CTGAGGCTCTCACTGAGAGTGAAGTGCAGCACTGAGAGAGAGCGCCGCGGAGAGAA 1380
 Qy 1827 TGCCCATTTGTCAGAGAAAGTGAAGTGCAGCACTGAGAGAGAGCGCCGCGGAGAGAA 1886
 Db 1381 TGCCCATTTGTCAGAGAAAGTGAAGTGCAGCACTGAGAGAGAGCGCCGCGGAGAGAA 1440
 Qy 1887 GCGGCGAGGCGCAAGTTGTCAGAAATGAGAGCGCTCTCCCAAGCGGCGCTCCAGAG 1946
 Db 1441 GCGGCGAGGCGCAAGTTGTCAGAAATGAGAGCGCTCTCCCAAGCGGCGCTCCAGAG 1500
 Qy 1947 AGGCTTCTTCCGATCTCTCTGACCAAGCGGCGCTTTTGAGAGCTCGTCTCTGATCTCGC 2006
 Db 1501 AGGCTTCTTCCGATCTCTCTGACCAAGCGGCGCTTTTGAGAGCTCGTCTCTGATCTCGC 1560
 Qy 2007 GCAAGAGTACACAGAGTGTTCAGAAAGCTCTTCTGAGGAGTGAAGGAGCGGTTTACGA 2066
 Db 1561 GCAAGAGTACACAGAGTGTTCAGAAAGCTCTTCTGAGGAGTGAAGGAGCGGTTTACGA 1620
 Qy 2067 GTGACTTGTGAGAGCTCTGAGGAGCGGAGCTGAGCGAGAACACCTCTGAGGCGCAAGCC 2126
 Db 1621 GTGACTTGTGAGAGCTCTGAGGAGCGGAGCTGAGCGAGAACACCTCTGAGGCGCAAGCC 1680
 Qy 2127 TGCGTCCAGAGCGAGCTCTCTTCTCCGCGGAGTGGTGGAGCCCGCTCACTAGAGAG 2186
 Db 1681 TGCGTCCAGAGCGAGCTCTCTTCTCCGCGGAGTGGTGGAGCCCGCTCACTAGAGAG 1727
 Qy 2187 TTGGGGAAGCTTTCACTTCTGAGAGATGCTCAACAACATGCTGGGACATCGAACCT 2246
 Db 1728 CTGAGAAAGTACT----- 1740
 Qy 2247 CCTACCTGTGATCCCTCTCTGCGGAGTGGAGGCTTACCCCGTCTTAGTCCGCGG 2306
 Db 1741 ----- 1740
 Qy 2307 AATGGAAGTGGGCGGCGGAGCGGAGCGGAGTGGGCGGAGCGGCGGAGCGG 2366
 Db 1741 AGTGGCAAGTGGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1765
 Qy 2367 GGCACCTGGAGCGGCGGAGTGTGAGAGAGGCTGAGAGCTGCTCCGCGGAGCTGTT 2426
 Db 1766 ----- 1765
 Qy 2427 AGGACCTTCCCTAGGCGCGGAGTGTCAAGACATGAGAGACTGCTTTTCTGGGCGCT 2486
 Db 1766 ----- 1765
 Qy 2487 GATGGAAGTGTCTATCCGCTGGGAGAGCGGAGCACTCCGCGCTCTCGAGAGCAC 2546
 Db 1766 ----- 1765
 Qy 2547 TGCGCTCAAGTGAAGTGGGCGGAGCGGAGCAAGAAATTGAAGCGGAAGTCTGGGAA 2606
 Db 1766 -----AA 1767

Qy 2607 GCTTAAAGTCTGAGAGGAGGAGCCCGGTTGAGAGAGAGAGCGGAACTCTTGGACAAAGC 2666
 Db 1768 GCTTAAAGTCTGAGAGGAGGAGCCCGGTTGAGAGAGAGAGCGGAACTCTTGGACAAAGC 1827
 Qy 2667 CTTGACAAAGCCAAAGAGTCCGCTCCGCGCGGAGTGGAGAGTGGCGGCGGCGGCG 2726
 Db 1828 CTTGACAAAGCCAAAGAGTCCGCTCCGCGCGGAGTGGAGAGTGGCGGCGGCGGCG 1887
 Qy 2727 GGGGCGGAGAGAGCTCAAGCTTCAAGACATATTTGCTCATTTTCTGGCAATTTCT 2786
 Db 1888 GGGGCGGAGAGAGCTCAAGCTTCAAGACATATTTGCTCATTTTCTGGCAATTTCT 1947
 Qy 2787 CAGAGTGAAGAAATAGTCAAGACCGAGAGTGAAGCGGAGGCTCGAGAGAGC 2846
 Db 1948 CAGAGTGAAGAAATAGTCAAGACCGAGAGTGAAGCGGAGGCTCGAGAGAGC 2007
 Qy 2847 GCACTTTTCTTCTGAAAAGTATATGAGGAGCTGAATGAGCTTCTGAGAGGTTGTT 2906
 Db 2008 GCACTTTTCTTCTGAAAAGTATATGAGGAGCTGAATGAGCTTCTGAGAGGTTGTT 2067
 Qy 2907 ACCGTTTTTATTTGACACAGAAAGAAAGTGTCTGCTCCCTTCGCGGAAATCTCT 2966
 Db 2068 ACCGTTTTTATTTGACACAGAAAGAAAGTGTCTGCTCCCTTCGCGGAAATCTCT 2127
 Qy 2967 CTTTAAAGCTGAAGTGTCTGAGAGTGTCTGAGAGTGTCTGAGAGTGTCTGAGAGTGT 3026
 Db 2128 CTTTAAAGCTGAAGTGTCTGAGAGTGTCTGAGAGTGTCTGAGAGTGTCTGAGAGTGT 2187
 Qy 3027 CTTTCTTTTGTGCGCTTCTTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAG 3086
 Db 2188 CTTTCTTTTGTGCGCTTCTTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAG 2247
 Qy 3087 CTGGGAGTGTGATAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGT 3146
 Db 2248 CTGGGAGTGTGATAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGT 2307
 Qy 3147 ACTTGAAGAGCTGTTTGAAGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGT 3206
 Db 2308 ACTTGAAGAGCTGTTTGAAGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGT 2367
 Qy 3207 TCAAG 3212
 Db 2368 TCAAG 2373

RESULT 5
 US-09-834-291-2
 ; Sequence 2, Application US/09834291
 ; Patent No. US20020042064A1

; GENERAL INFORMATION: Peter
 ; APPLICANT: Krammer, Peter
 ; APPLICANT: Muller-Schilling, Martina
 ; APPLICANT: Oren, Moshe
 ; TITLE OF INVENTION: p53 Binding Areas
 ; FILE REFERENCE: 4121-122
 ; CURRENT APPLICATION NUMBER: US/09/834,291
 ; PRIOR FILING DATE: 2001-08-21
 ; PRIOR APPLICATION NUMBER: PCT/DE99/03343
 ; PRIOR FILING DATE: 1998-10-18
 ; PRIOR APPLICATION NUMBER: DE 198 47 779.1
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 720
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; US-09-834-291-2

Query Match 22.2%; Score 713; DB 9; Length 720;
 Best Local Similarity 100.0%; Pred. No. 3.3e-186;
 Matches 713; Conservative 0; Mismatches 0; Gaps 0;

```

QY 2500 GATCCCGCTGGGCGAGCGGCGGCTCCGCGCTCTCGAGAGACCACTGGCTCCAGCTT 2559
DB 1 GATCCCGCTGGGCGAGCGGCGGCTCCGCGCTCTCGAGAGACCACTGGCTCCAGCTT 60
QY 2560 GAGGTGGGCGTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2619
DB 61 GAGGTGGGCGTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 2620 TGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2679
DB 121 TGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 2680 GCCAAAGGCTCCGCTCCGCGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2739
DB 181 GCCAAAGGCTCCGCTCCGCGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 2740 GAGCTTACAGCTTACAGAGACATATTTCTCTCTTGGAGAGAGAGAGAGAGAGAGAG 2799
DB 241 GAGCTTACAGCTTACAGAGACATATTTCTCTCTTGGAGAGAGAGAGAGAGAGAGAG 300
QY 2800 TAACTCAGCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2859
DB 301 TAACTCAGCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 2860 CTCGAAAAGATTATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2919
DB 361 CTCGAAAAGATTATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 2920 GTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2979
DB 421 GTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 2980 AGTGGCTGCTGAGTGGTGTATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 3039
DB 481 AGTGGCTGCTGAGTGGTGTATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 540
QY 3040 CCTTTCTAGCTGACATCCCATGAGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3099
DB 541 CCTTTCTAGCTGACATCCCATGAGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 3100 TACTGCTTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3159
DB 601 TACTGCTTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 3160 GTTTTGAAGAGTCCCTGCTCAGAGATGCCAGCTTGCAGATGCTTATCAAG 3212
DB 661 GTTTTGAAGAGTCCCTGCTCAGAGATGCCAGCTTGCAGATGCTTATCAAG 713

```

RESULT 6

```

US-09-925-299-187
; Sequence 187, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (803)

```

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (853)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-187

```

```

Query Match
Best Local Similarity 8.9%; Score 287.4; DB 9; Length 859;
Matches 294; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1954 CTTCATCTCTCTGACCAACCGGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 2013
DB 19 CGTCCGCCACGCGCTCCAGCCGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 78
QY 2014 TGACACAGAGTGTTCMAAGAGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 2073
DB 79 TGACACAGAGTGTTCMAAGAGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 138
QY 2074 GGTGAGAGCTTCAAGGAGGAGCACTGGACAGGAGACACCTTGAAGCCAGCTTGC 2133
DB 139 GGTGAGAGCTTCAAGGAGGAGCACTGGACAGGAGACACCTTGAAGCCAGCTTGC 198
QY 2134 CCAGGCGAGAGCTCTTCTTCCGCGGTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2193
DB 199 CCAGGCGAGAGCTCTTCTTCCGCGGTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 258
QY 2194 AGCTTTTCACTTGGAGAGATGCTCAACACCACTGCTGAGATCTGAGCTTCTTACT 2253
DB 259 AGCTTTTCACTTGGAGAGATGCTCAACACCACTGCTGAGATCTGAGCTTCTTACT 318
QY 2254 CTGGT 2258
DB 319 CTGGT 323

```

RESULT 7

```

US-09-925-299-187
; Sequence 187, Application US/09925299
; Patent No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (803)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (853)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-187

```

```

Query Match
Best Local Similarity 8.9%; Score 287.4; DB 11; Length 859;
Matches 294; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1954 CTTCATCTCTCTGACCAACCGGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 2013
DB 19 CGTCCGCCACGCGCTCCAGCCGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 78
QY 2014 TGACACAGAGTGTTCMAAGAGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 2073

```

[illegible]

SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 16
 LENGTH: 2534
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-949-713-16

Query Match
 Best Local Similarity 7.0%; Score 226; DB 9; Length 2534;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GAGCCTTCTGGGAGTGAAGGAAACCGGTTTACGAGTACTTGGCTGAGAGCTTCAAGGGGCG 2092
 Db 1 GAGCCTTCTGGGAGTGAAGGAAACCGGTTTACGAGTACTTGGCTGAGAGCTTCAAGGGGCG 60
 QY 2093 GGCACTGGACGGAACACACCTTGAAGCCGCTGCTGCTCCAGGCGGAGCTGCTCTT 2152
 Db 61 GGCACTGGACGGAACACACCTTGAAGCCGCTGCTGCTCCAGGCGGAGCTGCTCTT 120
 QY 2153 CTCGGCGGGTGTGTGACACCCGCTCACTACGAGTGTGGGAGAGCTCTTCACTTGGAGG 2212
 Db 121 CTCGGCGGGTGTGTGACACCCGCTCACTACGAGTGTGGGAGAGCTCTTCACTTGGAGG 180
 QY 2213 ATTGCTCAACACCAATGCTGGCATCTGAGACCTCTACTCTGCT 2258
 Db 181 ATTGCTCAACACCAATGCTGGCATCTGAGACCTCTACTCTGCT 226

RESULT 11
 US-09-884-987-1
 Sequence 1, Application US/09884987
 Patent No. US20020102653A1
 GENERAL INFORMATION:
 APPLICANT: NAGATA, Shigekazu et al
 TITLE OF INVENTION: DNA CODING FOR HUMAN CELL SURFACE ANTIGEN
 FILE REFERENCE: 0020-4872P
 CURRENT APPLICATION NUMBER: US/09/884,987
 CURRENT FILING DATE: 2001-06-21
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 1
 LENGTH: 2534
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: polyA_site
 LOCATION: (1831)..(1836)
 NAME/KEY: mat_peptide
 LOCATION: (243)..()
 NAME/KEY: sig_peptide
 LOCATION: (195)..(242)
 NAME/KEY: CDS
 LOCATION: (195)..(1199)
 NAME/KEY: polyA_site
 LOCATION: (2352)..(2357)
 NAME/KEY: polyA_site
 LOCATION: (2518)..(2532)
 US-09-884-987-1

Query Match
 Best Local Similarity 7.0%; Score 226; DB 10; Length 2534;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GAGCCTTCTGGGAGTGAAGGAAACCGGTTTACGAGTACTTGGCTGAGAGCTTCAAGGGGCG 2092
 Db 1 GAGCCTTCTGGGAGTGAAGGAAACCGGTTTACGAGTACTTGGCTGAGAGCTTCAAGGGGCG 60
 QY 2093 GGCACTGGACGGAACACACCTTGAAGCCGCTGCTGCTCCAGGCGGAGCTGCTCTT 2152
 Db 61 GGCACTGGACGGAACACACCTTGAAGCCGCTGCTGCTCCAGGCGGAGCTGCTCTT 120
 QY 2153 CTCGGCGGGTGTGTGACACCCGCTCACTACGAGTGTGGGAGAGCTCTTCACTTGGAGG 2212
 Db 121 CTCGGCGGGTGTGTGACACCCGCTCACTACGAGTGTGGGAGAGCTCTTCACTTGGAGG 180

Db 121 CTCGGCGGGTGTGTGACACCCGCTCACTACGAGTGTGGGAGAGCTCTTCACTTGGAGG 180
 QY 2213 ATTGCTCAACACCAATGCTGGCATCTGAGACCTCTACTCTGCT 2258
 Db 181 ATTGCTCAACACCAATGCTGGCATCTGAGACCTCTACTCTGCT 226

RESULT 12
 US-10-027-632-145670/c
 Sequence 145670, Application US/10027632
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827.129
 CURRENT APPLICATION NUMBER: US/10/027,632
 CURRENT FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/218,006
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/167,363
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: US 60/156,358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146,002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 145670
 LENGTH: 816
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-145670

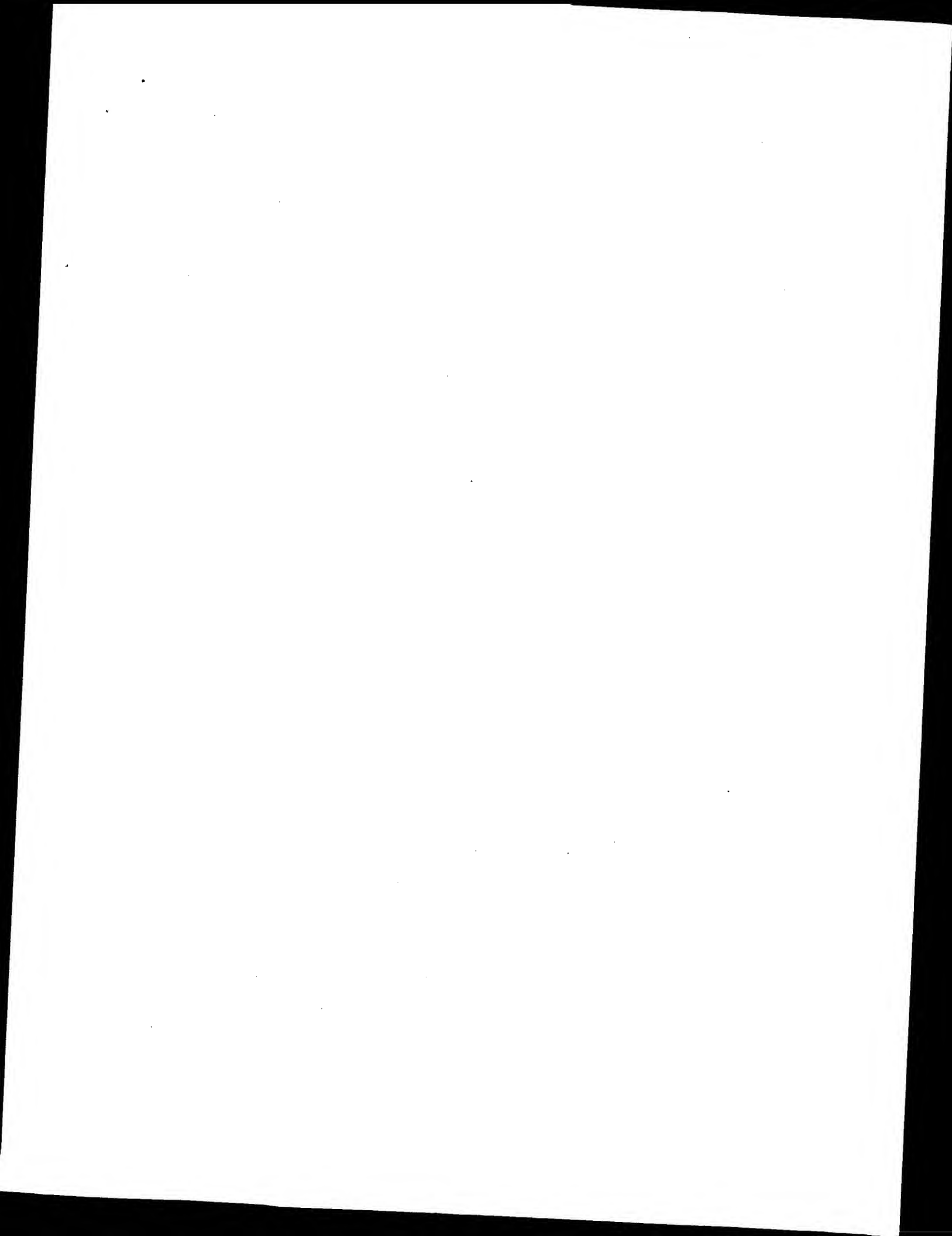
Query Match
 Best Local Similarity 1.9%; Score 60.8; DB 13; Length 816;
 Matches 83; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 697 TTTTATTTAAATGAACTTTTCACTTTTGAATAGTTTGAATTTCAAAAATTTGA 756
 Db 307 TGTTTTGTTTAAATGAACTTTTCACTTTTGAATAGTTTGAATTTCAAAAATTTGA 248
 QY 757 GAGATTAATACAGAGATGCCATATATACCATCTCTTATCCCACTTTCTTTGTGTAT 816
 Db 247 AAGATAGTACAGAAATATCCCAATATCTGACACACCGTTTCCCTGTTTTAACTTTAT 188

RESULT 13
 US-09-736-457-796/c
 Sequence 796, Application US/09736457
 Patent No. US2002016837A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Tongtong
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodges, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedrick, Tom
 APPLICANT: Carter, Darrick
 APPLICANT: Retter, Marc
 APPLICANT: Mannion, Jane
 APPLICANT: Fan, Liqun
 APPLICANT: Wang, Aijun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.478C15
 CURRENT APPLICATION NUMBER: US/09/736,457
 CURRENT FILING DATE: 2000-12-13
 NUMBER OF SEQ ID NOS: 1864

APPLICANT: McNeill, Patricia
APPLICANT:

Search completed: August 1, 2003, 13:36:32
Job time : 846.884 secs



Sun Aug 3 09:04:10 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model
August 1, 2003, 08:37:27 ; Search time 78.1573 seconds
(without alignments)

Run on: 10468.541 Million cell updates/sec

US-09-834-291-6

Title: 20
Sequence: 1 ggaagagccctgcaagcga 20

Scoring table: IDENTITY NJC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmb1.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ov.*
8: gb_ov.*
9: gb_ov.*
10: gb_ov.*
11: gb_ov.*
12: gb_ov.*
13: gb_ov.*
14: gb_ov.*
15: gb_ov.*
16: gb_ov.*
17: gb_ov.*
18: gb_ov.*
19: gb_ov.*
20: gb_ov.*
21: gb_ov.*
22: gb_ov.*
23: gb_ov.*
24: gb_ov.*
25: gb_ov.*
26: gb_ov.*
27: gb_ov.*
28: gb_ov.*
29: gb_ov.*
30: gb_ov.*
31: gb_ov.*
32: gb_ov.*
33: gb_ov.*
34: gb_ov.*
35: gb_ov.*
36: gb_ov.*
37: gb_ov.*
38: gb_ov.*
39: gb_ov.*
40: gb_ov.*
41: gb_ov.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	20	100.0	20	AX026094
2	18.4	92.0	20	AX026093
3	18.4	92.0	20	AX026095
4	18.4	92.0	20	AX026098
5	18.4	92.0	20	AX026112
6	18.4	92.0	266	AX026120
7	18.4	92.0	720	AX026090
8	18.4	92.0	2380	AX026091
9	18.4	92.0	2827	AX026089
10	18.4	92.0	3212	AX026089
11	18.4	92.0	45121	AX026089
12	18.4	92.0	187313	AX026089
13	18.4	92.0	167075	AX026089
14	18.4	92.0	714	AX026089
15	17.4	87.0	111702	AX026089
16	17.4	87.0	150493	AX026089
17	17.4	87.0	154687	AX026089
18	17.4	87.0	191676	AX026089
19	17.4	87.0	224401	AX026089
20	17.4	87.0	252630	AX026089
21	17.4	87.0	402	AX026089
22	16.8	84.0	1208	AX026089
23	16.8	84.0	1263	AX026089
24	16.8	84.0	1293	AX026089
25	16.8	84.0	1923	AX026089
26	16.8	84.0	1923	AX026089
27	16.8	84.0	2461	AX026089
28	16.8	84.0	11459	AX026089
29	16.8	84.0	29518	AX026089
30	16.8	84.0	70408	AX026089
31	16.8	84.0	83120	AX026089
32	16.8	84.0	85040	AX026089
33	16.8	84.0	85303	AX026089
34	16.8	84.0	92442	AX026089
35	16.8	84.0	93519	AX026089
36	16.8	84.0	93519	AX026089
37	16.8	84.0	93519	AX026089
38	16.8	84.0	108336	AX026089
39	16.8	84.0	108336	AX026089
40	16.8	84.0	110000	AX026089
41	16.8	84.0	110000	AX026089
42	16.8	84.0	110000	AX026089
43	16.8	84.0	110382	AX026089
44	16.8	84.0	116090	AX026089
45	16.8	84.0	116090	AX026089

ALIGNMENTS

RESULT 1
AX026094
LOCUS
DEFINITION
AX026094
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX026094 20 bp DNA linear PAT 16-SEP-2000
Sequence 6 from Patent DE19847779.
AX026094.1 GI:10187525
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Muller-Schilling, M., Kramer, P., and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: DE 19847779-C 6 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES location/Qualifiers
source 1 .20

BASE COUNT /mol_type="Homo sapiens"
ORIGIN /db_xref="taxon:9606"
8 a 6 c 5 g 1 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
Db 1 GGAAGAGCCCTGACAGCCA 20

RESULT 2

AX026093

LOCUS Sequence 5 from Patent DE19847779. 20 bp DNA linear PAT 16-SEP-2000
DEFINITION
ACCESSION AX026093
VERSION
KEYWORDS
SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Mueller-Schilling, M., Krammer, P. and Oren, M.
AUTHORS Novel receptor dna useful for identifying apoptosis-modulating
TITLE substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 5 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES location/Qualifiers
source 1 .20

BASE COUNT /organism="Homo sapiens"
ORIGIN /mol_type="genomic DNA"
7 a 7 c 5 g 1 t

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
Db 1 GGAAGAGCCCTGACAGCCA 20

RESULT 3

AX026095

LOCUS Sequence 7 from Patent DE19847779. 20 bp DNA linear PAT 16-SEP-2000
DEFINITION
ACCESSION AX026095
VERSION
KEYWORDS
SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Mueller-Schilling, M., Krammer, P. and Oren, M.
AUTHORS Novel receptor dna useful for identifying apoptosis-modulating
TITLE substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 7 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES location/Qualifiers
source 1 .20

/organism="Homo sapiens"

BASE COUNT /mol_type="genomic DNA"
ORIGIN /db_xref="taxon:9606"
9 a 5 c 5 g 1 t

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
Db 1 GGAAGAGCCCTGACAGCCA 20

RESULT 4

AX026098

LOCUS Sequence 10 from Patent DE19847779. 20 bp DNA linear PAT 16-SEP-2000
DEFINITION
ACCESSION AX026098
VERSION
KEYWORDS
SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Mueller-Schilling, M., Krammer, P. and Oren, M.
AUTHORS Novel receptor dna useful for identifying apoptosis-modulating
TITLE substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 10 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES location/Qualifiers
source 1 .20

BASE COUNT /organism="Homo sapiens"
ORIGIN /mol_type="genomic DNA"
7 a 7 c 5 g 1 t

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
Db 1 GGAAGAGCCCTGACAGCCA 20

RESULT 5

AX026112

LOCUS Sequence 24 from Patent DE19847779. 20 bp DNA linear PAT 16-SEP-2000
DEFINITION
ACCESSION AX026112
VERSION
KEYWORDS
SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Mueller-Schilling, M., Krammer, P. and Oren, M.
AUTHORS Novel receptor dna useful for identifying apoptosis-modulating
TITLE substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 24 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES location/Qualifiers
source 1 .20

BASE COUNT /organism="Homo sapiens"
ORIGIN /mol_type="genomic DNA"
7 a 7 c 5 g 1 t

Query Match 92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
1 GGACAGAGCCCTGACAGCCA 20

Db 1 GGACAGAGCCCTGACAGCCA 20

RESULT 6
LOCUS AX026120 266 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 32 from Patent DE19847779.
ACCESSION AX026120
VERSION AX026120.1 GI:10187551
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 32 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source 1. .266
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 49 a 72 c 110 g 35 t

ORIGIN
Query Match 92.0%; Score 18.4; DB 6; Length 266;
Best Local Similarity 95.0%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
160 GGACAGAGCCCTGACAGCCA 179

Db 160 GGACAGAGCCCTGACAGCCA 179

RESULT 7
LOCUS HSA011034 266 bp DNA linear PRI 20-JAN-1999
DEFINITION Homo sapiens DNA for enhancer of CD95 gene, partial.
ACCESSION AJ011034
VERSION AJ011034.1 GI:4165483
KEYWORDS CD95 gene; enhancer.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Mueller, M., Wilder, S., Bannasch, D., Israeli, D., Lehlbach, K.,
Li-Weber, M., Friedman, S.L., Galle, P.R., Stremmel, W., Oren, M. and
Kramer, P.H.
TITLE p53 activates the CD95 (Apo-1/Fas) gene in response to DNA damage
JOURNAL J. Exp. Med. 188 (11), 2033-2045 (1998)
MEDLINE 99059827
PUBMED 9841917

REFERENCE 2 (bases 1 to 266)
AUTHORS Mueller, M.
TITLE Direct Submmission
JOURNAL Submitted (08-SEP-1998) Mueller M., Department of Internal Medicine
IV, University Hospital, Bergheimerstr. 58, Heidelberg, 69115,
GERMANY
FEATURES
source 1. .266
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
gene 159. .178
/gene="CD95"
enhancer 159. .178
/gene="CD95"
/evidence="experimental"

BASE COUNT 49 a 72 c 110 g 35 t

ORIGIN
Query Match 92.0%; Score 18.4; DB 9; Length 266;
Best Local Similarity 95.0%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
160 GGACAGAGCCCTGACAGCCA 179

Db 160 GGACAGAGCCCTGACAGCCA 179

RESULT 8
LOCUS AX026090 720 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 2 from Patent DE19847779.
ACCESSION AX026090
VERSION AX026090.1 GI:10187521
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 2 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source 1. .720
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 142 a 181 c 216 g 181 t

ORIGIN
Query Match 92.0%; Score 18.4; DB 6; Length 720;
Best Local Similarity 95.0%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20
160 GGACAGAGCCCTGACAGCCA 179

Db 160 GGACAGAGCCCTGACAGCCA 179

RESULT 9
LOCUS AX026091 2380 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 3 from Patent DE19847779.
ACCESSION AX026091
VERSION AX026091.1 GI:10187522
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 3 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source 1. .2380
/organism="Homo sapiens"
/mol_type="genomic DNA"

BASE COUNT 579 a /db_xref="taxon:9606"
 ORIGIN 595 c 568 g 638 t

Query Match 92.0%; Score 18.4; DB 6; Length 2380;
 Best Local Similarity 95.0%; Pred. No. 2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20
 DB 1820 GGACAGCCCTGACAGCCA 1839

RESULT 10
 LOCUS AX026092 2827 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 4 from Patent DE19847779.
 ACCESSION AX026092
 VERSION AX026092.1 GI:10187523
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating
 JOURNAL substances potentially useful for cancer chemotherapy
 DEUTSCHES KREBSFORSCH (DE)
 Patent: DE 19847779-C 4 03-FEB-2000;
 FEATURES
 source Location/Qualifiers
 1.2827
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 2827;
 Best Local Similarity 95.0%; Pred. No. 1.9e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20
 DB 2267 GGACAGCCCTGACAGCCA 2286

RESULT 11
 LOCUS AX026089 3212 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 1 from Patent DE19847779.
 ACCESSION AX026089
 VERSION AX026089.1 GI:10187520
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating
 JOURNAL substances potentially useful for cancer chemotherapy
 DEUTSCHES KREBSFORSCH (DE)
 Patent: DE 19847779-C 1 03-FEB-2000;
 FEATURES
 source Location/Qualifiers
 1.3212
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 778 a 784 c 809 g 841 t
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 6; Length 3212;

Best Local Similarity 95.0%; Pred. No. 1.9e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20
 DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 12
 LOCUS AX695635 45121 bp DNA linear PAT 31-MAR-2003
 DEFINITION Sequence 1262 from Patent WO03008583.
 ACCESSION AX695635
 VERSION AX695635.1 GI:29418787
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Morris, D.W. and Engelhard, E.K.
 TITLE Novel compositions and methods for cancer
 JOURNAL Patent: WO 03008583-A 1262 30-JAN-2003;
 Sagres Discovery (US)
 FEATURES
 source Location/Qualifiers
 1.45121
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 45121;
 Best Local Similarity 95.0%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20
 DB 10653 GGACAGCCCTGACAGCCA 10672

RESULT 13
 LOCUS AL157394 187313 bp DNA linear PRI 22-AUG-2001
 DEFINITION Human DNA sequence from clone RPI1-399019 on chromosome 10,
 complete sequence.
 ACCESSION AL157394
 VERSION AL157394.15 GI:15384622
 KEYWORDS
 SOURCE HTG.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Blakey, S.
 TITLE Direct Submission
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Aug 31, 2001 this sequence version replaced gi:14161146.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL, Sw, SWISSPROT, Tr, TrEMBL, Wp, WORMPEP. Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr10

RP11-399019 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pSpace 6
This sequence is the entire insert of clone RP11-399019 The true left end of clone RP11-496423 is at 166408 in this sequence. The true right end of clone RP11-30415 is at 18704 in this sequence.

FEATURES

source

1. 187313

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="10"

/clone="RP11-399019"

/clone_1fb="RP11-11.2"

/clone_1fb="100156"

/note="Sequence confirmed by AC015461 sequenced by WIBR."

100157..100198

/note="Sequence from AC015461 sequenced by WIBR."

105808..105972

/note="Sequence from AC015461 sequenced by WIBR."

105973..105989

/note="Sequence confirmed by AC015461 sequenced by WIBR."

55669 a 36398 c 36888 g 58358 t

BASE COUNT

ORIGIN

Query Match

Best local Similarity 95.0%; Pred. No. 1e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGAAGAGCCTGACAGCCA 20

Db 144628 GGACAGAGCCTGACAGCCA 144647

RESULT 14

AC139121

LOCUS

DEFINITION

AC139121 167075 bp DNA linear HTG 05-FEB-2003

ordered pieces.

AC139121.2 GI:28209433

HTG; HTG_PHASE2; HTG_DRAFT.

Keywords: Didelphis virginiana (North American opossum)

Didelphis virginiana

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Marsupialia; Didelphimorphia; Didelphidae; Didelphis.

1 (bases 1 to 167075)

Ahter, N. Antropellis, A., Ayle, K., Beckstrom-Sternberg, S.M.,

Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooke, S.,

Carla, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,

Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,

Latic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B.,

Margulies, E.H., Mastello, C., Maskeri, B., McDowell, J.,

Paguitan, C., Pearson, R., Portnoy, M.E., Prasad, A.,

Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,

Stentrop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,

Wehrthy, K.D., Wiggins, L., Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 167075)

Green, E.D.

Direct Submission

Submitted (24-JAN-2003) NIH Intramural Sequencing Center, 8717

REFERENCE
3 (bases 1 to 167075)
Green, E.D.

TITLE
JOURNAL

Submitted (05-FEB-2003) NIH Intramural Sequencing Center, 8717

Submitted (05-FEB-2003) NIH Intramural Sequencing Center, 8717

On Feb 5, 2003 this sequence version replaced gi:27884867.

Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc_zoo@nigri.nih.gov

Project Information

Center project name: dvk

Center clone name: 039j13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; N/A; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 166451 bases at least Q40

Consensus quality: 166589 bases at least Q30

Consensus quality: 166648 bases at least Q20

Insert size: 150000; agarose-fp

Insert size: 166675; sum-of-contigs

Quality coverage: 10.81x in Q20 bases; sum-of-contigs

Quality coverage: 9.73x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 117048: contig of 117048 bp in length

* 117049 117148: gap of unknown length

* 117149 119135: contig of 1987 bp in length

* 119136 119235: gap of unknown length

* 119236 133059: contig of 13824 bp in length

* 133060 133159: gap of unknown length

* 133160 140947: contig of 7788 bp in length

* 140948 141047: gap of unknown length

* 141048 167075: contig of 26028 bp in length.

FEATURES

source

1. 167075

/organism="Didelphis virginiana"

/mol_type="genomic DNA"

/db_xref="taxon:9267"

/clone="LB3-39013"

/clone_1fb="LB3"

1. 117048

/note="assembly_fragment"

clone end: 17

vector side:left"

117149..119135

/note="assembly_fragment"

119236..133059

/note="assembly_fragment"

misc_feature 133160.140947
/note="assembly_fragment"
141048.167075
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

BASE COUNT 48094 a 31698 c 33625 g 53258 t 400 others
ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 167075;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGACCCCTGACAGC 18
|||||
Db 72852 GGAAGACCCCTGACAGC 72869

RESULT 15
BV077964 714 bp DNA linear STS 31-MAY-2003
LOCUS S212P60375FC6.T0 CZECHII/E1 Mus musculus STS genomic, sequence
DEFINITION BV077964
ACCESSION BV077964.1 GI:31193759
VERSION STS.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 714)
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,B.S., Lindblad-Toh,K. and Daly,M.J.
The mosaic structure of variation in the laboratory mouse genome
Nature 420 (6915), 574-578 (2002)
22354684
12466852

COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 714
Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SVimU, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGSVC3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.
location/Qualifiers

1. 714
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/E1"
/db_xref="taxon:10090"
/map="16 22-505 92651644-92652127"
/clone_11b="CZECHII/E1"
<1..>714

STS
BASE COUNT 202 a 153 c 160 g 199 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 11; Length 714;
Best Local Similarity 94.7%; Pred. No. 7.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGAAGACCCCTGACAGCCA 20
|||||
Db 248 GGAAGACCCCTGACATGCCA 266

Search completed: August 1, 2003, 17:32:44
Job time: 81.1573 secs

Sun Aug 3 09:04:15 2003

US-09-834-291-6.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds
(without alignments)
7098.748 Million cell updates/sec

Title: US-09-834-291-6
Perfect score: 20
Sequence: 1 ggaagaagccctgacagca 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18.4	92.0	793 12	BI763679 603049567
2	18.4	96.2	9	AL540709 AL540709
3	18.4	92.0	1089 13	BQ072834 AGENCOURT
4	17.4	87.0	255 13	BQ091744 UNM21B02

5	17.4	87.0	319 12	BI784680
6	17.4	87.0	693 10	BE272685
7	17.4	87.0	813 10	BE272685
8	17.4	87.0	1062 13	BQ070571
9	16.8	84.0	267 14	T95521
10	16.8	84.0	389 14	T95521
11	16.8	84.0	421 29	B2764825
12	16.8	84.0	426 9	AA233020
13	16.8	84.0	434 12	BM109280
14	16.8	84.0	440 13	BU962533
15	16.8	84.0	450 10	BE646301
16	16.8	84.0	454 14	T70290
17	16.8	84.0	455 10	BF175093
18	16.8	84.0	460 14	T87958
19	16.8	84.0	466 13	B0513664
20	16.8	84.0	490 28	BH028903
21	16.8	84.0	498 9	AV727532
22	16.8	84.0	508 28	BH029384
23	16.8	84.0	519 9	AI887588
24	16.8	84.0	521 10	BF962452
25	16.8	84.0	548 9	AA442710
26	16.8	84.0	554 9	AA490194
27	16.8	84.0	563 28	A2630360
28	16.8	84.0	569 13	BQ362540
29	16.8	84.0	577 12	B1924611
30	16.8	84.0	577 28	BH027622
31	16.8	84.0	587 10	AW950883
32	16.8	84.0	589 12	BM739855
33	16.8	84.0	590 9	AA625121
34	16.8	84.0	590 10	BE353461
35	16.8	84.0	591 13	BY470998
36	16.8	84.0	600 28	B2167939
37	16.8	84.0	605 13	BK474351
38	16.8	84.0	607 14	CB161114
39	16.8	84.0	612 28	BH900152
40	16.8	84.0	646 9	AV713997
41	16.8	84.0	659 9	AL877982
42	16.8	84.0	669 13	BQ512401
43	16.8	84.0	672 13	BU266199
44	16.8	84.0	675 9	AL864641
45	16.8	84.0	677 12	BM700586

ALIGNMENTS

RESULT 1
BI763679/c 793 bp mRNA linear EST 25-SEP-2001
LOCUS 603049567F1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5189752 5',
DEFINITION mRNA sequence.
ACCESSION BI763679
VERSION BI763679.1 GI:15755257
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 793)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: The I.M.A.G.E. Consortium (LNL)
CDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1M11474 row: e column: 17
High quality sequence stop: 786.

FEATURES

source

Location/Qualifiers
1. 793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5189752"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT 183 a 265 c 208 g 137 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 793;
Best Local Similarity 95.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GGAAGAGCCCTGACAGCCA 20
398 GGACAGCCCTGACAGCCA 379

RESULT 2

LOCUS

AL540709

DEFINITION

AL540709 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DB002YN18

ACCESSION

AL540709

VERSION

AL540709.2

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 962)

AUTHORS

Li W.B., Gruber C., Jessee J. and Polayes D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

COMMENT

On Feb 15, 2001 this sequence version replaced gi:12871113.

Contract: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Invitrogen. This sequence belongs to sequence cluster 5554.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DB002Y09QPI&cluster=5554.r. Contact :

Feng Liang Email: fliang@litech.com URL:

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID: CS0DB002Y09QPI.

Location/Qualifiers

1. 962

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DB002YN18"

/issue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime and enriched, the Not I and EcoRV sites of the pCMVSPORT 6 vector. into

Library was not normalized."

FEATURES

source

BASE COUNT 211 a 302 c 264 g 181 t 4 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 962;
Best Local Similarity 95.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GGAAGAGCCCTGACAGCCA 20
94 GGACAGCCCTGACAGCCA 75

RESULT 3

LOCUS

B0072834

DEFINITION

AGENCOURT 6763280 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5756324

ACCESSION

B0072834

VERSION

B0072834.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1089)

AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM12796 row: 1 column: 21

High quality sequence stop: 645.

Location/Qualifiers

1. 1089

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5756324"

/issue_type="Leukocyte"

/lab_host="DH10B"

/clone_lib="NIH_MGC_118"

/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of

non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.7 kb, insert size range

1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 027. Note:

this is a NIH MGC Library."

BASE COUNT 280 a 263 c 333 g 212 t 1 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 1089;
Best Local Similarity 95.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GGAAGAGCCCTGACAGCCA 20
369 GGACAGCCCTGACAGCCA 388

RESULT 4

LOCUS

B0091744

DEFINITION

WMN21B02 Canine Brain CDNA Library Canis familiaris CDNA 5', mRNA

255 bp

mRNA

linear EST 08-APR-2002

RESULT 6		693 bp	mRNA	linear	EST 13-JUL-2000
BE272685				clone IMAGE:3347822 5',	
LOCUS	BE272685				
DEFINITION	601105061F1 NIH_MGC_14 Homo sapiens CDNA				
	mRNA sequence.				
ACCESSION	BE272685				
VERSION	BE272685.1	GI:9147035			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 693)
AUTHORS NIH-MGC
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Cloned through the I.M.A.G.E. Consortium (LNL)
Plate: LICM39 row: 1 column: 15
High quality sequence stop: 609.
Location/Qualifiers

FEATURES

1..693
Location/Qualifiers

BASE COUNT

197 a 137 c 212 g 147 t

Query Match

Best Local Similarity 94.7%; Pred. No. 5.6e+02; Length 693;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GAAAGCCCTGACAGCCA 20
556 GAAAGCCCTGACAGCCA 574

RESULT 7

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BE296494 813 bp mRNA linear EST 20-JUL-2000
LOCUS BE296494
DEFINITION mRNA sequence.. NIH_MGC_17 Homo sapiens CDNA clone IMAGE:3530249 5',
ACCESSION BE296494
VERSION BE296494.1 GI:9180056
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 813)
AUTHORS NIH-MGC
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Cloned through the I.M.A.G.E. Consortium (LNL)
Plate: LICM202 row: 1 column: 18
High quality sequence stop: 536.
Location/Qualifiers

FEATURES

1..536
Location/Qualifiers

us-09-834-291-6.rst

source

1..813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3530249"
/issue_type="3530249"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7, Site 1: XhoI; Site 2: XhoI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 201 a 170 c 260 g 182 t

Query Match

Best Local Similarity 87.0%; Score 17.4; DB 10; Length 813;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GAAAGCCCTGACAGCCA 20
553 GAAAGCCCTGACAGCCA 571

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

B0070571 1062 bp mRNA linear EST 02-APR-2002
LOCUS B0070571
DEFINITION AGENCOURT_6839245 NIH_MGC_47 Homo sapiens CDNA clone IMAGE:5922971
ACCESSION B0070571
VERSION B0070571.1 GI:19899617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS NIH-MGC
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Cloned through the I.M.A.G.E. Consortium (LNL)
Plate: LICM090 row: 1 column: 12
High quality sequence stop: 624.
Location/Qualifiers

FEATURES

source

1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5922971"
/issue_type="5922971"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Brain; Vector: pOTB7, Site 1: XhoI; Site 2: XhoI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH-MGC Library.

BASE COUNT 197 a 352 c 317 g 196 t
 ORIGIN

Query Match 85.0%; Score 17; DB 13; Length 1062;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAAGAGCCCTGACAG 17
 |||||
 768 GGAAGAGCCCTGACAG 764

RESULT 9 267 bp mRNA linear EST 27-MAR-1995
 T95521
 LOCUS
 DEFINITION Y46G11.1 Soares fetal liver spleen INF5 Homo sapiens cDNA clone
 IMAGE:120788 5' similar to gb:M55905 MALATE OXIDOREDUCTASE (HUMAN
); mRNA sequence.

ACCESSION T95521
 VERSION T95521.1 GI:734145
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston
 R., Williamson, A., Wohldmann, P. and Wilson, R.
 The WashU-Merck EST Project

TITLE Unpublished
 JOURNAL
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1976
 High quality sequence stops: 113 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1976 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 113.
 Location/Qualifiers

FEATURES
 source
 1..267
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:473333"
 /db_xref="taxon:9606"
 /clone="IMAGE:120788"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="MDH10B (ampicillin resistant)"
 /clone_lib="Soares fetal liver spleen INF5"
 /note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5' AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pTZ19 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 72 a 54 c 57 g 76 t 8 others
 ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 267;
 Best Local Similarity 90.0%; Pred. No. 7.6e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGAAGAGCCCTGACAGCCA 20
 |||||

Db 136 GGAAGAGCCCTGACAGCCA 155
 |||||

RESULT 10 389 bp mRNA linear EST 15-MAR-1995
 T79683
 LOCUS
 DEFINITION Y471C09.1 Soares fetal liver spleen INF5 Homo sapiens cDNA clone
 IMAGE:113680 5' similar to gb:M55905 MALATE OXIDOREDUCTASE (HUMAN
); mRNA sequence.

ACCESSION T79683
 VERSION T79683.1 GI:698192
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston
 R., Williamson, A., Wohldmann, P. and Wilson, R.
 The WashU-Merck EST Project

TITLE Unpublished
 JOURNAL
 COMMENT Other_ESTS: y471C09.s1
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1814
 High quality sequence stops: 244 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1814 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 244.
 Location/Qualifiers

FEATURES
 source
 1..389
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:469297"
 /db_xref="taxon:9606"
 /clone="IMAGE:113680"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="MDH10B (ampicillin resistant)"
 /clone_lib="Soares fetal liver spleen INF5"
 /note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5' AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pTZ19 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 75 c 89 g 112 t 8 others
 ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 389;
 Best Local Similarity 90.0%; Pred. No. 8.9e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGAAGAGCCCTGACAGCCA 20
 |||||

Db 179 GGAAGAGCCCTGACAGCCA 198
 |||||

RESULT 11 421 bp DNA linear GSS 13-MAR-2003
 B2764825/c
 B2764825

DEFINITION SALK_127083.30.15.n Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_127083.30.15.n, genomic
survey sequence.
B2764825
B2764825..1 GI:28937378
GSS.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of Atg24160.
Class: TDNA tagged.

FEATURES
SOURCE

Location/Qualifiers
1..421
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_127083.30.15.n"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT
ORIGIN

113 a
77 c
87 g
144 t

Query Match 84.0%; Score 16.8; DB 29; Length 421;
Best Local Similarity 90.0%; Pred. No. 9,1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGAAAGCCCTGACAGCCA 20
| ||||| |||||
Db 395 GAAAAAGCCCTGAAAAAGCCA 376

RESULT 12
LOCUS
DEFINITION
VERSION
ACCESSION
KEYWORDS
SOURCE
ORGANISM

AA233020 426 bp mRNA linear EST 06-AUG-1997
z146a08.r1 Soares NHMPL SI Homo sapiens cDNA IMAGE:666422 5'
similar to gb:U55505 MALATE OXIDOREDUCTASE (HUMAN); mRNA sequence.
AA233020.1 GI:1856013
EST.

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 426)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Joet, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, T., Wyllie,
J., Waterston, R. and Wilson, R.
Washu-Merck EST Project 1997
Unpublished

COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Insert length: 1352 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 410.
Location/Qualifiers
1. .426
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5428454"
/db_xref="taxon:9606"
/clone="IMAGE:666422"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NHPu.S1"
/note="Organ: mixed (see below); Vector: pT773D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBW, pregnant uterus
2NBHP, and fetal heart 2NBH1.9W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 266232-265223,
340488-345479, and 484488-489479."

BASE COUNT
119 a . 90 c 96 g 121 t

ORIGIN

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 9; Length 426;
Matches 18; Conservativity 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GCAAAAGCCCTGACAGCA 20
|||||
Db 301 GCAAGGCCCTGACAGCA 320
|||||

RESULT 13
BM109280/c

LOCUS
BM109280
DEFINITION
EST556816 potato roots Solanum tuberosum cDNA clone cPRO3B10 5' end
/RNA sequence.
ACCESSION
BM109280
VERSION
BM109280.1 GI:117070044
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum
Bukarjoca; Vitididplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 434)
van der Hoeven, R., Sun, H., Karameyeva, S.A., Tsai, J., Van Aken, S.,
Tanksley, S., and Baker, B.
Tanksley, T., Chieningo, A., Bougri, O., Buell, C.R., Renning, C.,
Generation of ESTs from potato roots
Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
Location/Qualifiers

source

1. 434

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="CPR03B10"

/tissue_type="roots"

/dev_stage="in vitro grown stem cuttings"

/lab_host="SOLR"

/clone_1lb="potato roots"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after plating the stem cuttings from in vitro grown plants on medium."

BASE COUNT 125 a 98 c 82 g 129 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 434;
Best Local Similarity 90.0%; Pred. No. 9.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGAAGAGCCCTGACAGCCA 20
|||||
405 GGAAGAGCCGAGACAGCCA 386

RESULT 14
BU962533 440 bp mRNA linear EST 21-OCT-2002
LOCUS
DEFINITION AGENCOURT 10616779 NIH MGC 169 Mus musculus CDNA clone
IMAGE:6743287 5', mRNA sequence.
ACCESSION BU962533
VERSION BU962533.1 GI:24192105
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 440)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM3082 row: 1 column: 06
High quality sequence stop: 440.
Location/Qualifiers
1. 440
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6743287"
/lab_host="DH10B (T1-phage-resistant)"
/clone_1lb="NIH MGC 169"
/note="Organ: Testicles; Vector: pDNR-LIB; Site 1: SfiI (ggcgatcgcc); Site 2: SfiI (ggcgatcgcc); CDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAAGCAGATGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCCGAGCGCGCAGATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC library."

FEATURES
source

BASE COUNT 188 a 81 c 74 g 95 t 2 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 440;
Best Local Similarity 90.0%; Pred. No. 9.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGAAGAGCCCTGACAGCCA 20
|||||
8 GGAAGAGCCCTGACAGCCA 27

RESULT 15
BE646301 450 bp mRNA linear EST 05-SEP-2000
LOCUS
DEFINITION 7864d08.x1 NCI_CGAP_P128 Homo sapiens CDNA clone IMAGE:3289167 3', mRNA sequence.
ACCESSION BE646301
VERSION BE646301.1 GI:9970612
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 450)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap/
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmer-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.llnl.gov
Seq primer: -40UP from GIBCO.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3289167"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_P128"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP_P128 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 87 a 135 c 116 g 112 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 450;
Best Local Similarity 90.0%; Pred. No. 9.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGAAGAGCCCTGACAGCCA 20
|||||
433 GGAAGAGCCCTGACAGCCA 414

Search completed: August 1, 2003, 20:42:54
Job time : 72.4754 secs

Sun Aug 3 09:04:15 2003

us-09-834-291-6.rst

Page 8

Sun Aug 3 09:04:11 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 / Search time 6.69868 Seconds
(without alignments)
8059.612 Million cell updates/sec

Title: US-09-834-291-6

Perfect score: 1 ggaagagccctgacagca 20

Sequence:

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq, 190un03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	266	21	AAZ88700
2	16.8	84.0	402	22	AAZ88700
3	16.8	84.0	502	25	AAZ88700
4	16.8	84.0	1444	13	AAZ88700
5	16.8	84.0	1752	12	AAZ88700
6	16.8	84.0	1923	13	AAZ88700
7	16.8	84.0	1923	13	AAZ88700
8	16.8	84.0	1963	24	AAZ88700

US-09-834-291-6.rng

9	16.8	84.0	2390	22	AAZ88700
10	16.8	84.0	2463	22	AAZ88700
11	16.8	84.0	3085	22	AAZ88700
12	16.4	82.0	180	18	AAZ88700
13	16.4	82.0	936	24	AAZ88700
14	16.4	82.0	2214	21	AAZ88700
15	16.4	82.0	2214	21	AAZ88700
16	16.4	82.0	2214	25	AAZ88700
17	15.8	79.0	251	21	AAZ88700
18	15.8	79.0	345	21	AAZ88700
19	15.8	79.0	500	24	AAZ88700
20	15.8	79.0	624	23	AAZ88700
21	15.8	79.0	4057	17	AAZ88700
22	15.8	79.0	4796	22	AAZ88700
23	15.8	79.0	4796	22	AAZ88700
24	15.8	79.0	10519	23	AAZ88700
25	15.8	79.0	10519	23	AAZ88700
26	15.8	79.0	10519	23	AAZ88700
27	15.8	79.0	13337	22	AAZ88700
28	15.8	79.0	13337	23	AAZ88700
29	15.8	79.0	13337	23	AAZ88700
30	15.8	79.0	21129	22	AAZ88700
31	15.8	79.0	22700	22	AAZ88700
32	15.8	79.0	29255	23	AAZ88700
33	15.8	79.0	31813	22	AAZ88700
34	15.8	79.0	31813	22	AAZ88700
35	15.8	79.0	31813	22	AAZ88700
36	15.8	79.0	110608	24	AAZ88700
37	15.8	79.0	349980	24	AAZ88700
38	15.4	77.0	508	21	AAZ88700
39	15.4	77.0	508	22	AAZ88700
40	15.4	77.0	508	25	AAZ88700
41	15.4	77.0	1253	24	AAZ88700
42	15.4	77.0	3812	10	AAZ88700
43	15.4	77.0	465237	24	AAZ88700
44	15.4	77.0	465237	24	AAZ88700
45	15.4	77.0	465237	25	AAZ88700

ALIGNMENTS

RESULT 1
ID AAZ88700 standard: DNA; 266 BP.

AAZ88700; (first entry)

11-MAY-2000 (first entry)

Human CD95 receptor intron 1 fragment.

p53; CD95 receptor; human; screening; apoptosis-modulation;

cancer chemotherapy; ss.

Homo sapiens.

Location/Qualifiers

Key protein_bind

16-OCT-1998; 98DE-1047779.

16-OCT-1998; 98DE-1047779.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

Krammer P, Mueller-Schilling M, Oren M;

CDNA encoding nove
Novel protein kina
CDNA encoding nove
Lactococcal promot
Streptococcus poly
Human secreted pro
Human secreted pro
Secreted protein-e
Receptor #71 parti
Human ORX polynu
Human colon cancer
Human genome-deriv
Human prostate exp
Mouse neuron restir
Human nervous syst
Genomic sequence #
Genomic sequence #
Human nervous syst
Genomic sequence #
Genomic sequence #
Human DNA-binding
Human cardiovascular
Human immune/haema
Proionibacterium
Human ovarian and
Human reproductive
Human CDNA differe
Bifidobacterium lo
CDNA encoding huma
Colon tumor relat
Human lung specif
Human lung specif
Human oestrogen re
Gene encoding huma

XX WPI: 2000-162245/15.

XX Novel receptor DNA useful for identifying apoptosis-modulating
XX substances potentially useful for cancer chemotherapy
XX
XX Claim 2; Fig 4; 12pp; German.

XX This invention describes a novel p53-binding region of a human CD95
XX receptor DNA molecule. The p53-binding region, or a vector containing
XX it, can be used to screen for apoptosis-modulating substances
XX potentially useful for cancer chemotherapy. This sequence represents a
XX fragment of the human CD95 receptor intron 1 which contains a p53 binding
XX region described in the method of the invention.

XX Sequence 266 BP; 49 A; 72 C; 110 G; 35 T; 0 other;

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 21; Length 266;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCCA 20
Db 160 GGAAGAGCCCTGACAGCCCA 179

RESULT 2
AAFe4773
ID AAF64773 standard; cDNA; 402 BP.

AC AAF64773;
XX
XX 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 529.
XX
XX Human; cyclostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; ss.

XX Homo sapiens.
XX
XX WO200102568-A2.
XX
XX 11-JAN-2001.

XX
XX 30-JUN-2000; 2000WO-US18374.
XX
XX 02-JUL-1999; 99US-0142310.
XX
XX 02-JUL-1999; 99US-0142311.

XX (CHIR) CHIRON CORP.
XX (HSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
XX Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson J, Kassam A;
XX Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
XX Kita D, Garcia V, Jones LW, Strache-Crain B;
XX WPI: 2001-091805/10.

XX Library of polynucleotides for diagnosing a cancerous state of a
XX mammalian cell and detecting cancer, particularly of the colon or
XX prostate, comprises 3351 human polynucleotide sequences -
XX
XX Claim 9; Page 620; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human
XX polynucleotides. The library is used to detect differentially expressed
XX genes correlated with a cancerous state of a mammalian cell and can
XX detect colon, prostate, breast and lung cancer. The library can be used
XX to produce probes for detection of mRNA and to produce additional copies
XX of the polynucleotides. The probes can be used for chromosome mapping of
XX the polynucleotide and for detection of transcription levels. Ribozymes

CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 402 BP; 72 A; 121 C; 102 G; 106 T; 1 other;

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 22; Length 402;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCCA 20
Db 346 GGAAGAGCCCTGACAGCCCA 365

RESULT 3
ABZ19870
ID ABZ19870 standard; cDNA; 502 BP.

AC ABZ19870;
XX
XX 23-JAN-2003 (first entry)

XX Group III cDNA cancer related clone SEQ ID NO:2296.
XX
XX Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
XX immune response; virology; immunology; microbiology; molecular biology;
XX recombinant DNA technology; gene; ss.

XX Homo sapiens.
XX
XX WO200278516-A2.
XX
XX 10-OCT-2002.

XX
XX 28-MAR-2002; 2002WO-US10421.
XX
XX 30-MAR-2001; 2001US-280255P.
XX
XX 28-AUG-2001; 2001US-31563P.
XX
XX 09-JAN-2002; 2002US-347313P.

XX (CORI) CORIXA CORP.
XX
XX Wang T, Wang S, Bangur CS, Gaiger A;
XX WPI: 2003-058387/05.

XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
XX preventing and treating cancer expressing CT or CP mRNA antigens, and
XX recombinant DNA techniques -
XX
XX Claim 1; SEQ ID 2296; 207pp; English.

XX ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
XX ABP54446 to ABP54472 represent protein (II) sequences, from the present
XX invention. (I) and (II) have cytostatic activity and can be used in gene
XX therapy and vaccines. (I), (II), antibodies and compositions from the
XX cancer, which expresses CT or CP mRNA antigens. They are useful for
XX stimulating immune response. They can also be used in virology,
XX immunology, microbiology, molecular biology and recombinant DNA
XX techniques.

XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 502 BP; 131 A; 124 C; 163 G; 84 T; 0 other;

Sun Aug 3 09:04:11 2003

us-09-834-291-6.rng

Query Match 84.0%; Score 16.8; DB 25; Length 502;
Beet local similarity 90.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GGAAAGCCCTGACAGCCA 20
159 GGCAAGCCCTCAGAGCCA 178
Db
RESULT 4
AAS40992 standard; CDNA; 1444 BP.
ID AAS40992
AC AAS40992;
XX 17-DEC-2001 (first entry)
DT
XX CDNA encoding novel human enzyme polypeptide #208.
DE
XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
XX ligase; hyperproliferative disorder; immunodeficiency disorder;
XX autoimmune disorder; neurological disorder; metabolic disorder;
XX inflammatory disorder; cardiovascular disorder; reproductive disorder;
XX blood-related disorder; infectious disorder; gene therapy; cytostatic;
XX anti arthritic; nephrotoxic; anticoagulant; ss.
XX Homo sapiens.
OS WO20155301-A2.
XX
XX 02-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US01239.
PF
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-MAR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0214886.
XX 28-JUN-2000; 2000US-0215335.
XX 30-JUN-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216860.
XX 07-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217486.
XX 11-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 14-AUG-2000; 2000US-0224518.
XX 26-JUL-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225477.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227100.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240980.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246619.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancers

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU2814), and the cDNA and genomic sequences encoding them. The various polypeptides of the invention may comprise functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention may comprise the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), in the immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. hemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AA640785-AA641664 represent cDNA sequences encoding for the novel enzyme polypeptides of the invention.

(Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO accession 1447272, http://www.wipo.int/pub/published_pct_sequences).

CY March		84.0%	Score 16.0	DB 22;	Length 1444;
Local Similarity		90.0%	Pred. No.	1e+02;	
Conservative	0;	Mismatches	2;	Indels	0; Gaps
1 GCAGAAAGCCTGACACAACCA	20				
960 GCAGAGCCTGTGACACAACCCA	979				

ID	AAQ22695	standard; DNA; 1752 BP.
XX		
XX	AAQ22695;	
AC		
XX		
DT	25-MAR-2003	(updated)
DT	24-JUL-1992	(first entry)
XX		
DE		Sequence encoding mitochondrial NAD(P)+-dependent malate enzyme.
XX		
KW		Carbon metabolism; pyruvate formation; ss.
XX		
OS		Homo sapiens.
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1752
XX		/tag= a
XX		
EN	DE4028618-A.	
XX		
PD	12-MAR-1992.	
XX		
PF	08-SEP-1990;	90DE-4028618.
XX		
PR	08-SEP-1990;	90DE-4028618.
XX		
PR	19-JUN-1991;	91DE-4120178.
XX		
PA	(BOEH) BOEHRINGER INGELHEIM.	
XX		
PI	Dworkin MB, Leob , Leober G, Kryetek E, Maurefogy I;	
XX		
DR	WPI, 1992-089407/12.	
XX		
P	P-FSDB; AAR21845.	
XX		
PT		Human mitochondrial NAD(P)-dependent malate enzyme - used to
PT		study formation of pyruvate from aminoacid(s) in tumour cells
XX		
PS	Claim 2; Page 12-13; 20pp; German.	
XX		
CC		The inventors claim mitochondrial NAD(P)+-dependent malate enzyme
CC		and DNA encoding it. AAR21845 has 5' and 3' non-coding regions. The
CC		enzyme catalyses conversion of malate to pyruvate. Both the DNA and
CC		dividing cells, esp. pyruvate formation from amino acids.
CC		(Updated on 25-MAR-2003 to correct FA field.)
XX		
SO		Sequence 1752 BP; 562 A; 326 C; 386 G; 478 T; 0 other;
XX		
Query Match		84.0%; Score 16.8; DB 13; Length 1752;
Best Local Similarity		90.0%; Pred. No. 1e+02;
Matches 18; Conservative		0; Mismatches 2; Indels 0; Gaps 0;
OY		
	1	GGAAAGCCCTGACAGCCA 20
DB	1483	GCAAGCCCTGACAGCCA 1502
RESULT 6		
AAQ23258		
ID	AAQ23258	standard; DNA; 1923 BP.
XX		
AC	AAQ23258;	
XX		
DT	31-JUL-1992	(first entry)
XX		
DE		Mitochondrial NAD(P)+-dependent malate enzyme.
XX		
KW		C-metabolism; tumour; pyruvate; T-lymphocyte; ss.
XX		
OS		Homo sapiens.
XX		
Key		Location/Qualifiers
FT		
CDS		


```

XX  /tag= a
XX  /product= malate_enzyme
XX  90..149
XX  sig.peptide
XX  /tag= b
XX  150..1844
XX  mat.peptide
XX  /tag= c
XX  1905..1910
XX  polyA_signal
XX  /tag= d
XX  /note= "homologous to poly(A) signal AATAAA"
XX
XX  WC9204448-A.
XX  19-MAR-1992.
XX  23-AUG-1991.
XX  19-JUN-1991.
XX  08-SEP-1990.
XX  91MO-EP01602.
XX  91DE-4120178.
XX  90DE-4028618.
XX
XX  (BOEH ) BOEHRINGER INGELHEIM.
XX  Dworkin MB, Loeber G, Krystek E, Maurerfог I, Frubels B;
XX  WPI: 1992-114355/14.
XX  P-PSDB; AAR23356.
XX
XX  New human mitochondrial malate enzyme and DNA encoding it - for
XX  studying carbon metabolism in cells, also specific antibodies for
XX  purification and assay
XX
XX  Claim 1; Page 46 + Fig 3; 60pp; German.
XX
XX  The sequence may be used to study C-metabolism in rapidly dividing
XX  (tumor) cells, esp. pyruvate formation from amino acids
XX  CC The enzyme was first isolated from the supernatant of mitochondria
XX  CC preparations from the transformed human T-lymphocyte cell line
XX  CC 1301. tryptic fragments were partially sequenced and used as a
XX  CC basis for the design of oligonucleotides. These were used in PCR
XX  CC for amplification of malate enzyme encoding DNA in a cDNA bank
XX  CC prep. from fibrosarcoma Hs913. Amplified fragments were subcloned
XX  CC in pUC18, sequenced and used to probe the fibrosarcoma bank.
XX  CC A 1923bp insert was isolated and cloned in Bluescript KS+.
XX  CC The poly(A+) tail is not included in this sequence.
XX
XX  Sequence 1923 BP; 598 A; 373 C; 436 G; 516 T; 0 other;
XX
XX  Query Match 84.0%; Score 16.8; DB 13; Length 1923;
XX  Best Local Similarity 90.0%; Pred. No. 1e+02;
XX  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX  1 GGAAGAGCCCTGACAGCA 20
XX  1572 GCAAGGCCCCGACAGCA 1591
XX
XX  RESULT 7
XX  ABL67216
XX  ID ABL67216 standard; DNA; 1923 BP.
XX
XX  ABL67216;
XX  15-MAY-2002 (first entry)
XX
XX  Thyroid cancer related gene sequence SEQ ID NO:5553.
XX  Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX  stomach; lung; prostate; carcinoma; antitumor; cancerous;
XX  cytostatic; gene therapy; antineoplastic; Wilms' tumor; adenocarcinoma;
XX  gene; ds.
XX  Homo sapiens.
XX  WO200194629-A2.

```

```

XX  13-DEC-2001.
XX  30-MAY-2001; 2001WO-US10838.
XX
XX  05-JUN-2000; 2000US-209473P.
XX  05-JUN-2000; 2000US-209531P.
XX  18-SEP-2000; 2000US-233133P.
XX  18-SEP-2000; 2000US-23317P.
XX  20-SEP-2000; 2000US-234009P.
XX  20-SEP-2000; 2000US-234034P.
XX  20-SEP-2000; 2000US-234052P.
XX  20-SEP-2000; 2000US-234509P.
XX  22-SEP-2000; 2000US-234567P.
XX  22-SEP-2000; 2000US-234923P.
XX  25-SEP-2000; 2000US-235077P.
XX  25-SEP-2000; 2000US-235082P.
XX  25-SEP-2000; 2000US-235134P.
XX  25-SEP-2000; 2000US-235637P.
XX  26-SEP-2000; 2000US-235638P.
XX  26-SEP-2000; 2000US-235711P.
XX  27-SEP-2000; 2000US-235720P.
XX  27-SEP-2000; 2000US-235840P.
XX  27-SEP-2000; 2000US-235843P.
XX  28-SEP-2000; 2000US-236032P.
XX  28-SEP-2000; 2000US-236034P.
XX  28-SEP-2000; 2000US-236109P.
XX  28-SEP-2000; 2000US-236111P.
XX  28-SEP-2000; 2000US-236842P.
XX  29-SEP-2000; 2000US-236891P.
XX  29-SEP-2000; 2000US-237172P.
XX  02-OCT-2000; 2000US-237173P.
XX  02-OCT-2000; 2000US-237278P.
XX  02-OCT-2000; 2000US-237295P.
XX  02-OCT-2000; 2000US-237316P.
XX  02-OCT-2000; 2000US-237425P.
XX  03-OCT-2000; 2000US-237588P.
XX  03-OCT-2000; 2000US-237604P.
XX  03-OCT-2000; 2000US-237606P.
XX  03-OCT-2000; 2000US-237608P.
XX  01-NOV-2000; 2000US-244867P.
XX  01-NOV-2000; 2000US-245084P.
XX
XX  (AVAIL-) AVAILON PHARM.
XX  Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX  Sopet DR, Weaver Z;
XX  WPI: 2002-188264/24.
XX
XX  Screening for anti-neoplastic agent involves exposing cells to a
XX  chemical agent to be tested for anti-neoplastic activity, and
XX  determining a change in expression of a gene of a signature gene set
XX
XX  Claim 1; SEQ ID 5553; 44pp; English.
XX
XX  The present invention describes a method (M1) for screening for an
XX  anti-neoplastic agent. The method involves exposing cells to a chemical
XX  agent to be tested for anti-neoplastic activity, determining a change in
XX  expression of at least one gene (I) of a signature gene set where (I)
XX  comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX  CC ABL61664) or is at least 95% identical to (S), where a change in
XX  expression is indicative of anti-neoplastic activity. (I) has cytosolic
XX  activity and can be used in gene therapy. M1 can be used for screening
XX  an anti-neoplastic agent, and can be used for producing a product which
XX  is the data collected with respect to the anti-neoplastic agent as a
XX  result of M1, and the data is sufficient to convey the chemical
XX  structure and/or properties of the agent. M1 can be used in the

```

CC Treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX Sequence 1923 BP; 599 A; 373 C; 436 G; 515 T; 0 other.

Query Match 84.0%; Score 16.8; DB 24; Length 1923;
 Best Local Similarity 90.0%; Pred. No. 1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DY 1 GGAAGAGCCCTGACAGCCA 20
 DB 1572 GCAAGGCCCTGACAGCCA 1591

RESULT 8
 ABL89871
 ID ABL89871 standard; cDNA; 1963 BP.
 AC ABL89871;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 433.

Cytostatic; immunosuppressive; neurotropic; neuroproliferative; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antifungal;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 OS Homo sapiens.

XX WO200190304-A2.
 XX
 PD 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US16450.
 XX
 PF 19-MAY-2000; 2000US-205515P.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PA Birse CE, Rosen CA;
 XX
 PI WPI; 2002-122018/16.
 DR P-PSDB; ABB89462.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 XX prevention of neutral, immune system, muscular, reproductive, and
 XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 XX disorders.

Claim 4, SEQ ID NO 433; 2081bp + Sequence Listing; English.
 The invention relates to novel genes (ABL89449-ABL90853) and proteins
 (ABB89040-ABB89044) useful for preventing, treating or ameliorating
 medical conditions e.g. by protein or gene therapy. The genes are
 isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, Crohn's
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 Note: The sequence data for this patent did not form part of the

CC Printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1963 BP; 622 A; 365 C; 394 G; 581 T; 1 other.

Query Match 84.0%; Score 16.8; DB 24; Length 1963;
 Best Local Similarity 90.0%; Pred. No. 1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DY 1 GGAAGAGCCCTGACAGCCA 20
 DB 960 GCAAGGCCCTGACAGCCA 979

RESULT 9
 AAS40991
 ID AAS40991 standard; cDNA; 2390 BP.
 AC AAS40991;
 XX
 DT 17-DEC-2001 (first entry)
 XX
 DE cDNA encoding novel human enzyme polypeptide #207.

Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
 KW anti arthritic; neurotropic; anticoagulant; ss.

OS Homo sapiens.
 XX WO200155301-A2.
 XX
 PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01239.
 XX
 PF 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216447.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.

Sun Aug 3 09:04:11 2003

us-09-834-291-6.rng

PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 14-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 21-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 27-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 29-SEP-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239337.
 PR 13-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 XX
 XX
 PI MPI, 2001-465566/50.
 XX P-PsDB; AU23120.
 DR
 XX
 XX Novel polypeptides and polynucleotides useful for diagnosing,
 PT preventing, treating neural, immune system, muscular, reproductive,
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 PT diseases
 XX
 XX Claim 4; SEQ ID NO 216; 1180bp; English.
 PS
 XX The present invention relates to the isolation of novel human enzyme
 CC polypeptides (AU22915-AU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious diseases (e.g. influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AU240785-AU241684 represent cDNA sequences encoding for the novel human
 CC enzyme polypeptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 XX Sequence 3085 BP; 948 A; 579 C; 652 G; 905 T; 1 other;
 SQ
 Query Match 84.0%; Score 16.8; DB 22; Length 3085;
 Best Local Similarity 90.0%; Pred. No. 1 le+02;
 Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GGAAGGCTGACACCA 20
 DB 1668 GCAAGGCTGACACCA 1687
 ID AAT91285 standard; DNA; 180 BP.
 AC AAT91285;
 XX 27-APR-1998 (first entry)
 DE Lactococcal promoter p23.
 XX Neutral protease; npr gene; enzyme engineering; protein engineering;
 KM lactic acid bacterium; food; milk; cheese; ripening;
 KW Bacillus subtilis; promoter; ss.
 XX

XX human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antisthma;
 KW antiproliferatic; cardiac; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; ss.

Sun Aug 3 09:04:13 2003

US-09-834-291-6.rmp

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds
(without alignments)
795.924 Million cell updates/sec

Title: US-09-834-291-6
Perfect score: 20 ggaagccctgacagcca 20
Sequence: 1 ggaagccctgacagcca 20
Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	US-09-834-291-6	Sequence 6, Appli
2	18.4	92.0	20	US-09-834-291-5	Sequence 7, Appli
3	18.4	92.0	20	US-09-834-291-7	Sequence 10, Appli
4	18.4	92.0	20	US-09-834-291-10	Sequence 24, Appli
5	18.4	92.0	20	US-09-834-291-24	Sequence 32, Appli
6	18.4	92.0	266	US-09-834-291-2	Sequence 2, Appli
7	18.4	92.0	720	US-09-834-291-2	Sequence 3, Appli
8	18.4	92.0	2380	US-09-834-291-2	Sequence 4, Appli
9	18.4	92.0	2827	US-09-834-291-1	Sequence 1, Appli
10	18.4	92.0	3212	US-09-834-291-1	Sequence 250, App
11	16.4	84.0	1923	US-09-964-824-250	Sequence 134797,
12	16.4	82.0	824	US-10-027-632-134797	Sequence 1, Appli
13	16.4	82.0	9025608	US-10-156-761-1	Sequence 10989, A
14	15.8	79.0	438	US-09-918-995-10989	Sequence 15724, A
15	15.8	79.0	500	US-09-866-108-15724	Sequence 137408,
16	15.8	79.0	606	US-10-027-632-137408	Sequence 137408,

17	15.8	79.0	606	13	US-10-027-632-137409	Sequence 137409,
18	15.8	79.0	814	13	US-10-027-632-8488	Sequence 8488, Ap
19	15.8	79.0	907	13	US-10-027-632-137413	Sequence 137413,
20	15.8	79.0	907	13	US-10-027-632-137414	Sequence 137414,
21	15.8	79.0	4796	10	US-09-764-847-1627	Sequence 1627, Ap
22	15.8	79.0	4796	10	US-09-764-847-1628	Sequence 1628, Ap
23	15.8	79.0	10519	10	US-09-764-847-1628	Sequence 1628, Ap
24	15.8	79.0	10519	10	US-09-764-847-1628	Sequence 1629, Ap
25	15.8	79.0	13337	10	US-09-764-847-1629	Sequence 1629, Ap
26	15.8	79.0	13337	10	US-09-764-847-1629	Sequence 1629, Ap
27	15.8	79.0	13337	10	US-10-091-483-112	Sequence 1734, Ap
28	15.8	79.0	13337	10	US-10-092-154-1629	Sequence 1734, Ap
29	15.8	79.0	21129	9	US-09-764-891-1734	Sequence 922, App
30	15.8	79.0	31813	14	US-10-025-428-932	Sequence 922, App
31	15.8	79.0	31813	14	US-10-025-428-932	Sequence 366, App
32	15.8	79.0	508	10	US-09-833-263-366	Sequence 366, App
33	15.4	77.0	508	10	US-10-025-380-366	Sequence 111067,
34	15.4	77.0	654	13	US-10-027-632-111067	Sequence 111067,
35	15.4	77.0	654	13	US-10-027-632-111067	Sequence 111067,
36	15.4	77.0	654	13	US-10-027-632-111067	Sequence 111067,
37	15.4	77.0	789	14	US-10-198-846-10892	Sequence 10892, A
38	15.4	77.0	1119	14	US-10-001-857-68	Sequence 4234, Ap
39	15.4	77.0	1253	13	US-10-231-426-3	Sequence 68, Appli
40	15.4	77.0	18540	14	US-10-103-637-1	Sequence 3, Appli
41	15.4	77.0	90442	14	US-10-103-637-1	Sequence 1, Appli
42	15.4	77.0	465237	10	US-09-933-267A-1	Sequence 8, Appli
43	15.4	77.0	465237	10	US-09-834-291-6	Sequence 9, Appli
44	15.2	76.0	20	9	US-09-834-291-9	
45	15.2	76.0	20	9	US-09-834-291-9	

ALIGNMENTS

RESULT 1
US-09-834-291-6 Application US/09834291
Sequence 6, Appli
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-6
Query Match 100.0%, Score 20, DB 9, Length 20;
Best Local Similarity 100.0%, Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GGAAGCCCTGACAGCCA 20
1 GGAAGCCCTGACAGCCA 20
RESULT 2
US-09-834-291-5 Application US/09834291
Sequence 5, Appli
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-5

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 1 GGAAGAGCCCTGACAGCCA 20

RESULT 3
US-09-834-291-7
Sequence 7, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-7

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 1 GGAAGAGCCCTGACAGCCA 20

RESULT 4
US-09-834-291-10
Sequence 10, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-10

PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-10

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 1 GGAAGAGCCCTGACAGCCA 20

RESULT 5
US-09-834-291-24
Sequence 24, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-24

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 1 GGAAGAGCCCTGACAGCCA 20

RESULT 6
US-09-834-291-32
Sequence 32, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 266
US-09-834-291-32

Sun Aug 3 09:04:13 2003

us-09-834-291-6.inpb

TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-32

Query Match 92.0%; Score 18.4; DB 9; Length 266;
Best Local Similarity 95.0%; Pred. No. 9.8;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 160 GGACAGCCCTGACAGCCA 179

RESULT 7
US-09-834-291-2
Sequence 2, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 720
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-2

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 160 GGACAGCCCTGACAGCCA 179

Query Match 92.0%; Score 18.4; DB 9; Length 720;
Best Local Similarity 95.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 160 GGACAGCCCTGACAGCCA 179

RESULT 8
US-09-834-291-3
Sequence 3, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 2380
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-3

Query Match 92.0%; Score 18.4; DB 9; Length 2380;
Best Local Similarity 95.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 1820 GGACAGCCCTGACAGCCA 1839

RESULT 9
US-09-834-291-4
Sequence 4, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 2827
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-4

Query Match 92.0%; Score 18.4; DB 9; Length 2827;
Best Local Similarity 95.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 2267 GGACAGCCCTGACAGCCA 2286

RESULT 10
US-09-834-291-1
Sequence 1, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 3212
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-1

Query Match 92.0%; Score 18.4; DB 9; Length 3212;
Best Local Similarity 95.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20
DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 11

US-09-964-824A-250
; Sequence 250, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horriken, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 250
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-250

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 1923;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGAAGCCCTGACAGCCA 20
DB 1572 GCAAGCCCTGACAGCCA 1591

RESULT 12
US-10-027-632-134797/c
; Sequence 134797, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108627,129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134797
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134797

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 13; Length 824;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 AAAAGCCCTGACAGCCA 20
DB 694 AAAAGCCCTGACAGCCA 677

RESULT 13
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMTURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 14; Length 9025608;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GAAAGCCCTGACAGCCA 19
DB 8367754 GAAAGCCCTGAAAGCC 8367757

RESULT 14
US-09-918-995-10989
; Sequence 10989, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: FROM VARIOUS CDNA LIBRARIES
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/918, 995
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10989
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-10989

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 11; Length 438;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GAAAGCCCTGACAGCCA 20
DB 383 GAAAGCCCTGACAGCCA 401

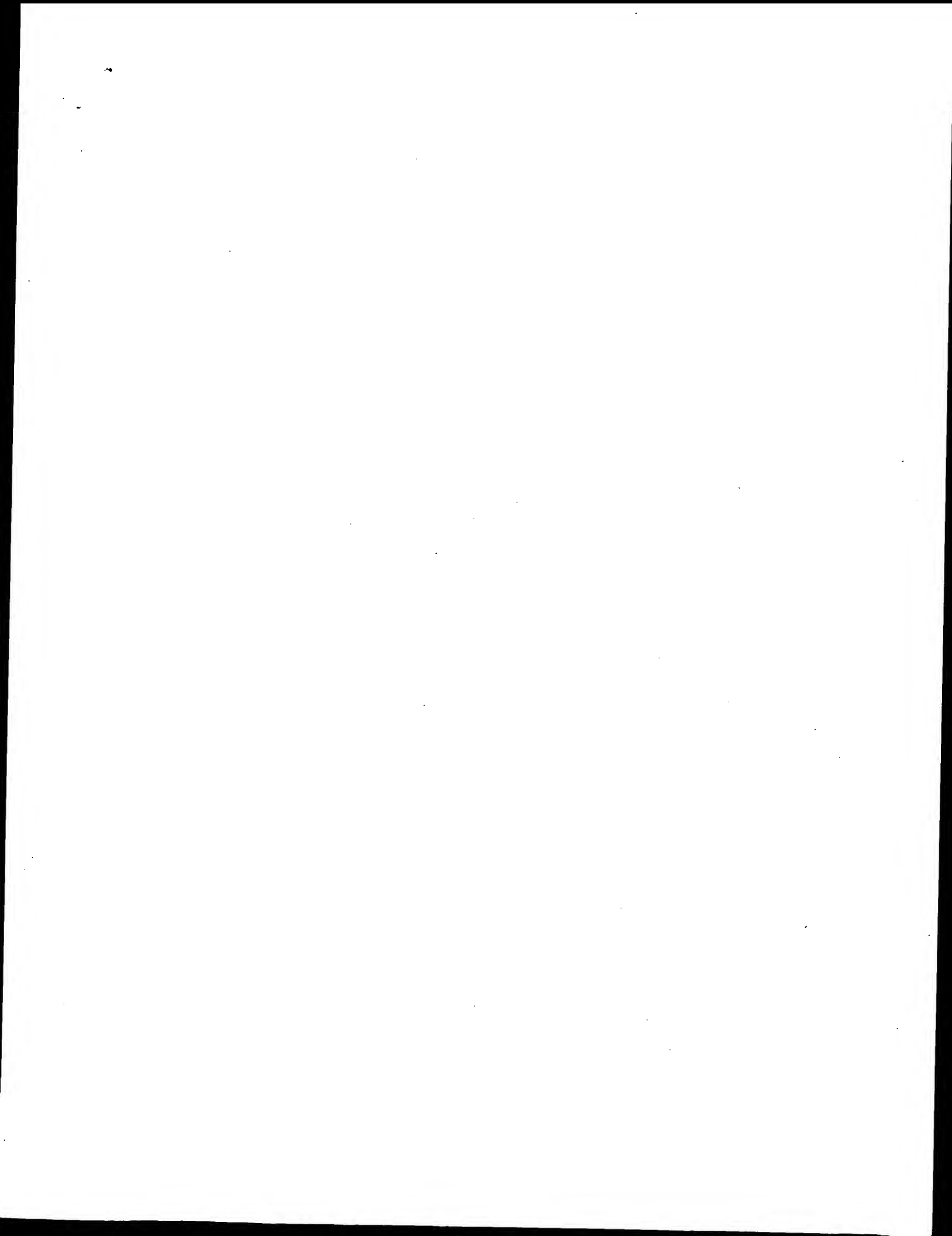
RESULT 15
US-09-866-108-15724

Sequence 15724, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharon G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/266,860
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 15752
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 15724
LENGTH: 500
TYPE: DNA
ORGANISM: Homo sapiens
US-09-866-108-15724

Query Match 79.0%; Score 15.8; DB 9;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Length 500;

QY 1 GGAAAGCCCTGACAGCC 19
DB 383 GGAAAGCCCTGACAGGC 401

Search completed: August 1, 2003, 13:36:44
Job time: 17.1861 secs



Sun Aug 3 09:04:12 2003

us-09-834-291-6.rml

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.959 Million cell updates/sec

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.959 Million cell updates/sec

Title: US-09-834-291-6
Perfect score: 20
Sequence: 1 ggaagagccctgacagcca 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 563978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.8	84.0	1923	4 US-09-220-132-85	Sequence 85, Appl
2	15.8	80.0	2214	4 US-08-489-847-113	Sequence 113, Appl
3	15.8	79.0	72	3 US-09-216-531-93	Sequence 93, Appl
4	15.8	79.0	4057	3 US-08-894-997-49	Sequence 49, Appl
5	15.2	76.0	738	4 US-09-252-991A-14234	Sequence 14234, A
6	15.2	76.0	1107	4 US-09-252-991A-14146	Sequence 14146, A
7	15.2	76.0	1521	4 US-09-252-991A-13990	Sequence 13990, A
8	15.2	76.0	1872	4 US-09-232-200-88	Sequence 88, Appl
9	15.2	76.0	1872	4 US-09-232-201-88	Sequence 88, Appl
10	15.2	76.0	1872	4 US-09-232-201-88	Sequence 88, Appl
11	15.2	76.0	4403765	3 US-09-103-840A-1	Sequence 2, Appl
12	15.2	76.0	4411529	3 US-09-103-840A-1	Sequence 17, Appl
13	14.8	74.0	2221	4 US-09-634-238-137	Sequence 1, Appl
14	14.8	72.0	1194	3 US-09-135-782-1	Sequence 1, Appl
15	14.4	72.0	1237	1 US-08-240-372-1	Sequence 1, Appl
16	14.4	72.0	1344	3 US-09-305-901-1	Sequence 1, Appl
17	14.4	72.0	1820	3 US-09-530-812A-1	Sequence 221, App
18	14.4	72.0	2470	4 US-08-965-762-13	Sequence 13, Appl
19	14.4	72.0	3800	4 US-09-911-927-13	Sequence 13, Appl
20	14.4	72.0	3800	4 US-09-911-927-15	Sequence 15, Appl
21	14.4	72.0	3800	4 US-09-911-882-13	Sequence 13, Appl
22	14.4	72.0	3800	4 US-09-911-882-15	Sequence 15, Appl
23	14.4	72.0	3800	4 US-09-911-888-13	Sequence 13, Appl
24	14.4	72.0	3800	4 US-09-911-888-15	Sequence 15, Appl
25	14.4	72.0	6516	4 US-08-861-527-105	Sequence 105, Appl
26	14.4	72.0	17138	3 US-09-813-819-3	Sequence 3, Appl
27	14.4	72.0	17138	3 US-09-813-819-3	Sequence 3, Appl

28	14.4	72.0	17138	4 US-09-920-048-3	Sequence 3, Appl
29	14.4	72.0	36181	4 US-08-311-711A-120	Sequence 120, Appl
30	14.4	71.0	90	2 US-08-155-171B-3	Sequence 3, Appl
31	14.2	71.0	90	2 US-08-435-998-3	Sequence 86, Appl
32	14.2	71.0	218	3 US-09-188-930-86	Sequence 2050, Ap
33	14.2	71.0	218	4 US-09-312-283C-86	Sequence 25, Appl
34	14.2	71.0	315	4 US-08-799-464A-25	Sequence 18, Appl
35	14.2	71.0	384	5 PCT-US95-09927-25	Sequence 15, Appl
36	14.2	71.0	387	3 US-08-855-531D-15	Sequence 25, Appl
37	14.2	71.0	387	3 US-08-855-531D-15	Sequence 15, Appl
38	14.2	71.0	387	3 US-08-855-531D-15	Sequence 25, Appl
39	14.2	71.0	387	3 PCT-US95-10904-25	Sequence 5, Appl
40	14.2	71.0	478	4 US-09-585-173B-5	Sequence 2613, Ap
41	14.2	71.0	600	4 US-09-252-991A-2613	Sequence 1267, Ap
42	14.2	71.0	669	4 US-09-328-352-1267	Sequence 9, Appl
43	14.2	71.0	769	2 US-08-538-711A-9	Sequence 9, Appl
44	14.2	71.0	769	3 US-08-725-027-9	Sequence 9, Appl
45	14.2	71.0	769	3 US-08-725-027-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-220-132-85
Sequence 85, Application US/09220132
Patent No. 6506607
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION OF PROSTATE CANCER
TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 1998-03-25
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 85
LENGTH: 1923
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-85

Query Match 84.0%; Score 16.8; DB 4; Length 1923;
Best Local Similarity 90.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20
DB 1572 GCAAGGCGCTGACAGCCA 1591
RESULT 2
US-09-489-847-113
Sequence 113, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
PRIOR APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12

EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
NUMBER OF SEQ ID NOS: 376
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 113

LENGTH: 2214

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (289)

OTHER INFORMATION: n equals a,t,g, or c

US-09-489-847-113

Query Match
Best Local Similarity 80.0%; Score 16; DB 4; Length 2214;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGCCCTGACAGCA 20
DB 305 GGGARAGGCTGACAGCA 324

RESULT 3
US-09-276-531-93
Sequence 93, Application US/09276531
Patent No. 6183968

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Lal, Preeti

APPLICANT: Hillman, Jennifer L.

APPLICANT: Yue, Henry

APPLICANT: Reddy, Roopa

APPLICANT: Guegler, Karl J.

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING

NUMBER OF SEQUENCES: 134

CORRESPONDENCE ADDRESS: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION

ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/276,531

FILING DATE: Herewith

CLASSIFICATION:

Prior Application DATA:

APPLICATION NUMBER: 60/079,677

FILING DATE: March 27, 1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Lynn E. Murry, Ph.D.

REGISTRATION NUMBER: 42,918

TELECOMMUNICATION INFORMATION: PA-0008 US

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 93:

SEQUENCE CHARACTERISTICS:

LENGTH: 72 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: BRSTU108
CLONE: 1729463
US-09-276-531-93

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 3; Length 72;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAAGCCCTGACAGCA 20
DB 32 GGAAGCCCTGACAGCA 50

RESULT 4
US-08-894-997-49

Sequence 49, Application US/08894997A
Patent No. 6270990

GENERAL INFORMATION:

APPLICANT: Anderson, David J

APPLICANT: Schoenheit, Christopher J

TITLE OF INVENTION: NEURON-RESTRICTIVE SILENCER FACTOR

FILE REFERENCE: 17810-502 NRSP

CURRENT APPLICATION NUMBER: US/08/894,997A

EARLIER FILING DATE: 1998-01-06

EARLIER APPLICATION NUMBER: PCT/US96/02817

EARLIER FILING DATE: 1996-03-01

EARLIER APPLICATION NUMBER: 08/398,590

NUMBER OF SEQ ID NOS: 35

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 49

LENGTH: 4057

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: gene

LOCATION: (1)..(4057)

OTHER INFORMATION: Human NSRP

US-08-894-997-49

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 3; Length 4057;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAAGCCCTGACAGCA 20
DB 1895 GGAAGCCCTGACAGCA 1913

RESULT 5
US-09-252-991A-14234
Sequence 14234, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

EARLIER FILING DATE: 1999-02-18

EARLIER APPLICATION NUMBER: 60/074,788

EARLIER FILING DATE: 1998-02-18

EARLIER APPLICATION NUMBER: US 60/094,190

NUMBER OF SEQ ID NOS: 33142.

SEQ ID NO 14234

LENGTH: 738

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-14234

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 4; Length 738;

Best Local Similarity 85.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAGGCTGACAGCCA 20
160 GGAAGGCTGACAGCCA 179

RESULT 6
US-09-252-991A-14146
Sequence 14146, Application US/09252991A

PATENT INFORMATION:
PATENT NO. 6551795
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 14146
LENGTH: 1107
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14146

Query Match 76.0%; Score 15.2; DB 4; Length 1107;
Best Local Similarity 85.0%; Pred. No. 72;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAGGCTGACAGCCA 20
597 GGAAGGCTGACAGCCA 616

RESULT 7
US-09-252-991A-13990/c
Sequence 13990, Application US/09252991A

PATENT INFORMATION:
PATENT NO. 6551795
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13990
LENGTH: 1521
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13990

Query Match 76.0%; Score 15.2; DB 4; Length 1521;
Best Local Similarity 85.0%; Pred. No. 77;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAGGCTGACAGCCA 20
739 GGAAGGCTGACAGCCA 720

RESULT 8
US-09-232-200-88/c
Sequence 88, Application US/09232200A

PATENT NO. 6288213
GENERAL INFORMATION:

APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATY ACID TRANSPORT PROTEINS
FILE REFERENCE: WH197-2193MB
CURRENT FILING DATE: 1999-01-14
PRIOR FILING DATE: 1998-01-15
PRIOR APPLICATION NUMBER: 60/093,491
PRIOR FILING DATE: 1998-07-20
PRIOR APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 88
LENGTH: 1872
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-232-200-88

Query Match 76.0%; Score 15.2; DB 3; Length 1872;
Best Local Similarity 85.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAGGCTGACAGCCA 20
46 GGAAGGCTGACAGCCA 27

RESULT 9
US-09-232-197-88/c
Sequence 88, Application US/09232197A

PATENT NO. 6300096
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATY ACID TRANSPORT PROTEINS
FILE REFERENCE: WH197-2193MB
CURRENT FILING DATE: 1999-01-14
PRIOR FILING DATE: 1998-01-15
PRIOR APPLICATION NUMBER: 60/093,491
PRIOR FILING DATE: 1998-07-20
PRIOR APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 88
LENGTH: 1872
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-232-197-88

Query Match 76.0%; Score 15.2; DB 4; Length 1872;
Best Local Similarity 85.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAGGCTGACAGCCA 20
46 GGAAGGCTGACAGCCA 27

RESULT 10
US-09-232-201-88/c

```

; Sequence 88, Application US/09232201A
; Patent No. 6348321
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21P3MC
; CURRENT APPLICATION NUMBER: US/09/232, 201A
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 88
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-232-201-88

```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 4; Length 1872;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGCA 20
DB 46 GGAAGATCTTGACAGCA 27

```

```

RESULT 11
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294338
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "a" bases at various positions throughout the sequence
US-09-103-840A-2

```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 3; Length 4403765;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGCA 20
DB 4172848 GGAAGAGCCCTGATATATCA 4172867

```

```

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A

```

```

; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 3; Length 4411529;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGCA 20
DB 4180600 GGAAGAGCCCTGATATATCA 4180619

```

```

RESULT 13
US-09-634-238-137
; Sequence 137, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glem, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137
; LENGTH: 2221
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2221)
; OTHER INFORMATION: n = A,T,C or G
US-09-634-238-137

```

```

Query Match
Best Local Similarity 74.0%; Score 14.8; DB 4; Length 2221;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGC 18
DB 283 GGAAGAGCCCTGACAGC 300

```

```

RESULT 14
US-09-135-782-1/C
; Sequence 1, Application US/09135782

```


Sun Aug 3 09:04:12 2003

us-09-834-291-6.rn1

Page 5

Patent No. 6027929
GENERAL INFORMATION:
APPLICANT: Xu, Shuang-yong
TITLE OF INVENTION: Method For Cloning And Producing The Nsp1 Restriction
TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The
TITLE OF INVENTION: Recombinant Nsp1 Restriction Endonuclease
FILE REFERENCE: NEB-143
CURRENT APPLICATION NUMBER: US/09/135,782
CURRENT FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 1
LENGTH: 1194
TYPE: DNA
ORGANISM: No. 6027929LOC SP.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1191)
US-09-135-782-1

Query Match 72.0%; Score 14.4; DB 3; Length 1194;
Best Local Similarity 93.8%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAAAGCCCTGACAG 17
DB 422 GAAAGCCCTGACAG 407

RESULT 15
US-08-240-372-1
Sequence 1, Application US/08240372
Patent No. 5741665
GENERAL INFORMATION:
APPLICANT: KATO, ELIE K.
APPLICANT: STUART, W. DORSEY
TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION
TITLE OF INVENTION: OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,372
FILING DATE: 10-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 3918-0003.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1237 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-240-372-1

Query Match 72.0%; Score 14.4; DB 1; Length 1237;
Best Local Similarity 93.8%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGCCCTGACAGCCA 20
DB 365 AAGCCCTGACAGACA 380

Search completed: August 1, 2003, 08:37:08
Job time: 13.6507 secs

Sun Aug 3 09:03:40 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds
(without alignments)
10468.541 Million cell updates/sec

Title: US-09-834-291-10
Perfect score: 1 ggaacagccctgacacgca 20
Sequence: IDENTITY_NUC
Scoring table: Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenBank:1:
1: gb ba:1:
2: gb hg:1:
3: gb in:1:
4: gb om:1:
5: gb ov:1:
6: gb pa:1:
7: gb ph:1:
8: gb pl:1:
9: gb pr:1:
10: gb ro:1:
11: gb st:1:
12: gb sy:1:
13: gb un:1:
14: gb vi:1:
15: em ba:1:
16: em fun:1:
17: em hum:1:
18: em in:1:
19: em mu:1:
20: em om:1:
21: em or:1:
22: em ov:1:
23: em pa:1:
24: em ph:1:
25: em pl:1:
26: em ro:1:
27: em st:1:
28: em un:1:
29: em vi:1:
30: em hg_hum:1:
31: em hg_inv:1:
32: em hg_other:1:
33: em hg_mus:1:
34: em hg_pla:1:
35: em hg_rod:1:
36: em hg_man:1:
37: em hg_vir:1:
38: em by:1:
39: em htgo_hum:1:
40: em htgo_mus:1:
41: em htgo_other:1:

Pred. No. is the number of results predicted by chance to have a

US-09-834-291-10.rge

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	20	100.0	20	6	AX026093
2	20	100.0	20	6	AX026098
3	20	100.0	20	6	AX026112
4	20	100.0	266	6	AX026120
5	20	100.0	266	6	HS4011034
6	20	100.0	720	6	AX026091
7	20	100.0	2380	6	AX026092
8	20	100.0	2827	6	AX026089
9	20	100.0	3212	6	AX026089
10	20	100.0	45121	6	AX695635
11	20	100.0	187313	9	AL157394
12	18.4	92.0	20	6	AX026094
13	18.4	92.0	167412	2	AL356352
14	18.4	92.0	192657	10	AL672160
15	18.4	92.0	193352	2	EX294664
16	18.4	92.0	199776	9	AC011499
17	18.4	92.0	204012	10	AL732613
18	18	90.0	81200	2	AC022086
19	18	90.0	145667	2	AL662817
20	18	90.0	214440	2	AC091422
21	18	90.0	215973	8	BF444228
22	17.4	87.0	380	2	AL339577
23	17.4	87.0	80961	2	AL929559
24	17.4	87.0	163933	5	AL589763
25	17.4	87.0	166772	9	AL587943
26	17.4	87.0	175770	2	AL365495
27	17.4	87.0	185049	9	AL365495
28	17.4	87.0	186325	2	AC074244
29	17.4	87.0	189485	2	AC141268
30	17.4	87.0	193347	2	AC125884
31	17.4	87.0	203246	5	AL954133
32	17.4	87.0	206373	2	AL155292
33	17.4	87.0	223831	2	AC116393
34	17.4	87.0	228659	2	AC126053
35	17.4	87.0	230097	2	AC099140
36	17.4	87.0	238103	2	AC135751
37	17.4	87.0	258131	2	AC108322
38	17.4	87.0	289818	2	AC112477
39	17	85.0	226723	2	AC126871
40	17	85.0	240214	2	AC105569
41	17	85.0	257842	2	AC128975
42	17	85.0	262681	2	AC121541
43	17	85.0	280469	2	AX026095
44	16.8	84.0	20	6	AX026097
45	16.8	84.0	20	6	AX026097

ALIGNMENTS

RESULT 1
LOCUS AX026093 20 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 5 from Patent DE19847779.
ACCESSION AX026093
VERSION AX026093.1 GI:10187524
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

Query Match	100.0%	Score 20;	DB 6;	Length 20;
Best Local Similarity	100.0%	Pred. No. 68;		
Matches	20;	Conservative		

QY	Indels	Gaps
1 GGACAGGCTCGACAGCCA 20	0;	0;
1 GGACAGGCTCGACAGCCA 20	0;	0;
1 GGACAGGCTCGACAGCCA 20	0;	0;

RESULT 2	20 bp	DNA	linear
AX026098			
LOCUS			
DEFINITION			
ACCESSION	Sequence 10 from Patent DE19847779.		
VERSION	AX026098		
FEATURES	AX026098.1	GI:10187529	
CURCD			
1			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Mueller-Schilling M., Kramer, P. and Oren, M.
 Novel receptor dna useful for identifying apoptosis-modulating
 substances potentially useful for cancer chemotherapy
 Patent: DE 19847779-C 10 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)

```

source      location/Qualifiers
1. .20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
7 a         7 c         5 g         1

```

```

Query Match      100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No 68;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGACAGACCCCTTACAAGCCA 20
        |||||
Db       1 GGACAGACCCCTTACAAGCCA 20

```

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Eukaryotes
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 Euteleostomi; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
Möller-Schilling, M., Krammer, P., and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
Substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 24 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
Location/Qualifiers
1..20
/organism="Homo sapiens"

Query Match	100.0%	Score 20;	DB 6;	Length 20;
Best Local Similarity	100.0%	Pred. No. 68;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.

Db

1 GGACAAAGCCCTGACAAAGCCA 20
|||||
1 GGACAAAGCCCTGACAAAGCCA 20

RESULT 4						
AX026120	AX026120					
LOCUS						
DEFINITION	AX026120					
ACCESSION	Sequence 32 from Patent DE19847779.	266 bp	DNA			
VERSION	AX026120					linear
KEYWORDS	AX026120.1	G1:10187551				PAT 16-SEP-2000
SOURCE	*					
ORGANISM	Homo sapiens (human)					
	Homo sapiens					

REFERENCE AUTHORS TITLE JOURNAL FEATURES SOURCE	1. Location/Qualifiers
1. <i>Chlamydia</i> ; <i>Eutheria</i> ; <i>Primates</i> ; <i>Catarrhini</i> ; <i>Hominiidae</i> ; <i>Homo</i> . Mueller-Schilling, M., Kramer, P. and Oren, M. Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy Patent: DE 1984779-C 32 03-FEB-2000; DEUTSCHES KREBSFORSCH (DE)	Location/Qualifiers 1. .266

BASE COUNT	ORIGIN
49 a	72 c 110 g 35 t

[illegible]

RESULT 5					
HSA011034	HSA011034	266 bp	DNA	linear	PRI 20-JAN-1999
LOCUS					
DEFINITION	Homo sapiens DNA for enhancer of CD95 gene, partial.				
VERSION	AJ011034				
KEYWORDS	AJ011034.1 GI:4165483				
SOURCE	CD95 gene; enhancer.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

REFERENCE	Mullerand; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Muller M., Wilder, S., Bannasch, D., Israeli, D., Lebach, K., Li Weber, M., Friedman, S.L., Galle, P.R., Stemmel, W., Oren, M. and Kramer, P.H.
TITLE	p53 anticancer drugs CD95 (Apo-1/Fas) gene in response to DNA damage
JOURNAL	J. Exp. Med. 188 (11), 2033-2045 (1998)
PUBMED	96059827

RECEIVED 9841917
 REFERENCE 2 (Pages 1 to 266)
 AUTHORS Mueller M.
 TITLE Direct Submission
 JOURNAL Submitted (08-SEP-1998) Mueller M., Department of Internal Medicine
 IV, University Hospital, Bergheimerstr. 58, Heidelberg, 69115,

FEATURES GERMANY Location/Qualifiers
 source 1. .266 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene 159. .178 /gene="CD95"
 159. .178 /gene="CD95"
 enhancer /evidence="experimental"
 BASE COUNT 49 a 72 c 110 g 35 t
 ORIGIN
 Query Match 100.0%; Score 20; DB 9; Length 266;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 160 GGACAGCCCTGACAGCCA 179
 QY 1 GGACAGCCCTGACAGCCA 20
 LOCUS AX026090 720 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 2 from Patent DE19847779.
 ACCESSION AX026090
 VERSION AX026090.1 GI:10187521
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating
 substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 19847779-C 2 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)
 FEATURES Location/Qualifiers
 source 1. .720 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 142 a 181 c 216 g 181 t
 ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 720;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGACAGCCCTGACAGCCA 20
 LOCUS AX026091 2380 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 3 from Patent DE19847779.
 ACCESSION AX026091
 VERSION AX026091.1 GI:10187522
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating
 substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 19847779-C 3 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)

FEATURES DEUTSCHES KREBSFORSCH (DE) Location/Qualifiers
 source 1. .2380 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 579 a 595 c 568 g 638 t
 ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 2380;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1820 GGACAGCCCTGACAGCCA 1839
 QY 1 GGACAGCCCTGACAGCCA 20
 LOCUS AX026092 2827 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 4 from Patent DE19847779.
 ACCESSION AX026092
 VERSION AX026092.1 GI:10187523
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating
 substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 19847779-C 4 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)
 FEATURES Location/Qualifiers
 source 1. .2827 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 728 a 676 c 657 g 766 t
 ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 2827;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGACAGCCCTGACAGCCA 20
 LOCUS AX026089 3212 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 1 from Patent DE19847779.
 ACCESSION AX026089
 VERSION AX026089.1 GI:10187520
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating
 substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 19847779-C 1 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)
 FEATURES Location/Qualifiers
 source 1. .3212 /organism="Homo sapiens"
 /mol_type="genomic DNA"

Query Match	100.0%;	Score 20;	DB 6;	Length 3212;
Best Local Similarity	100.0%;	Pred. No. 24;		
Matches 20;	Conservative 0;	Mismatches 0;		

Qy	1	GGACAAGCCCTGACAAGCCA	20
Db			
	2659	GGACAAGCCCTGACAAGCCA	2678

RESULT 10	AX695635	LOCUS	AX695635	45121 bp	DNA	linear	PAT 31-MAR-2003
		DEFINITION	Sequence 1262 from Patent WO03008563.				
		ACCESSION	AX695635				
		VERSION	AX695635.1				
		KEYWORDS	GI:29418787				
		SOURCE					
		ORGANISM	Homo sapiens (human)				
			Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

1 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Morris, D.W. and Engelhard, E.K.
Novel compositions and methods for cancer
patent: WO 03008563-A 1262 30-JAN-2003,
Sages Discovery (US)
Location/Qualifiers
1. .45121

BASE COUNT	13226	a	8836	c	9010	g	14049	t
ORIGIN	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"							

Query Match	100.0%;	Score 20;	DB 6;	Length 45121;
Best Local Similarity	100.0%;	Pred. No. 14;		
Matches 20;	Conservative 0;	Mismatches 0;		

y 1 GGACAAGCCCTGACAAGCCA 20
 |||||
 Db 10653 GGACAAGCCCTGACAAGCCA 10672

RESULT 11	LOCUS	DEFINITION	ACCESSION
AL157394	187113 bp	Human DNA sequence from clone RP11-399019 on chromosome 10, complete sequence.	AL157394

VERSION	AL157394.15	GI:15384622
KEYWORDS	HTG.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

1 (bases 1 to 187313)
Plakow

JOURNAL

COMMENT

Where differences assembly data is compared from overlapping clones. Together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate

Chemistry or covered by high quality data (i.e., phred quality >= 30) ; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em1, EMBL; SW1, SWISSPROT; Tr1, TREMBL; Wp1, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10> RP11-39019 is from the library RPc1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.choi.org/bacpac/home.htm> VECTOR: pBACe3.6

This sequence is the entire insert of clone RP11-39019 The true left end of clone RP11-496H3 is at 166408 in this sequence. The true right end of clone RP11-30415 is at 18704 in this sequence.

location/Qualifiers

1..187313

```
misc_feature      100119..100156
                    /note="Sequence confirmed by AC015461 sequenced by WIBR."
                    /db_xref="taxon:9606"
                    /chromosome="10"
                    /clone_lib="RPCT-11.2"
                    /clone_id="RP11-399019"
misc_feature      100157..100198
                    /note="Sequence confirmed by AC015461 sequenced by WIBR."
                    /note="Sequence from AC015461 sequenced by WIBR."
                    /note="Sequence from AC015461 sequenced by WIBR."
                    /note="Sequence confirmed by AC015461 sequenced by WIBR."
misc_feature      105808..105972
                    /note="Sequence confirmed by AC015461 sequenced by WIBR."
                    /note="Sequence confirmed by AC015461 sequenced by WIBR."
misc_feature      105973..105989
                    /note="Sequence confirmed by AC015461 sequenced by WIBR."
BASE COUNT       55669 a 36398 c 36888 g 58358 t
ORIGIN
```

Query Match	100.0%;	Score 20;	DB 9;	Length 187313;
Best Local Similarity	100.0%;	Pred. No. 11;		
Matches 20;	Conservative 0;	Mismatches 0;		

Qy 1 GGACACAGCCTTGACACAGCCA 20
|||
Db 144628 GGACACAGCCTTGACACAGCCA 144647

RESULT 12					
AX026094	AX026094	20 bp	DNA	linear	PAT 16-SEP-2000
LOCUS	Sequence 6 from Patent DE19847779.				
DEFINITION					
ACCESSION	AX026094				
VERSION	AX026094.1	GI:10187525			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Mammalia; Eutheria; Primates; Carnivora; Euteleostomi
Müller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 1984779-C 6 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

SOURCE

BASE COUNT	ORIGIN
8 a	6 c
	5 g
	1 t

Sun Aug 3 09:03:40 2003

us-09-834-291-10.rge

Query Match 92.0%; Score 18.4; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGACAGCCCTGACAGCCCA 20
1 CGAAGAGCCCTGACAGCCCA 20

RESULT 13
AL356352 167412 bp DNA linear HTG 21-OCT-2001
LOCUS Homo sapiens chromosome 1 clone RP5-859H16, *** SEQUENCING IN
DEFINITION PROGRESS *** 20 unordered pieces.
ACCESSION AL356352 GI:16304915
VERSION AL356352.8
KEYWORDS HTG: HTGS PHASE1; HTGS_CANCELED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS McLay, K.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Oct 21, 2001 this sequence version replaced gi:12539659.
Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: dj859H16
Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads Chemistry:
Chemistry: Dye-terminator ET-amersham; 26% of reads Chemistry:
Dye-terminator Big Dye; 73% of reads
Consensus quality: 15949 bases at least Q40
Consensus quality: 162148 bases at least Q30
Consensus quality: 163926 bases at least Q20
Insert size: 16512; sum-of-contigs
Insert size: 206683; 10.2% error; agarose-fp
Quality coverage: 4.08x in Q20 bases; sum-of-contigs Quality
coverage: 3.39x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 12019: contig of 12019 bp in length
* 12020 12119: gap of 100 bp
* 12120 32657: contig of 20536 bp in length
* 32658 32757: gap of 100 bp
* 32758 36105: contig of 3348 bp in length
* 36106 36205: gap of 100 bp
* 36206 41069: contig of 4864 bp in length
* 41070 41169: gap of 100 bp
* 41170 51754: contig of 10485 bp in length
* 51755 62397: gap of 100 bp
* 62398 73595: contig of 10643 bp in length
* 73596 83277: gap of 100 bp
* 83278 83377: contig of 9582 bp in length
* 83378 83377: gap of 100 bp

FEATURES
source

83378 98845: contig of 15468 bp in length
* 83378 98845: gap of 100 bp
* 98846 101173: contig of 2228 bp in length
* 98846 101173: gap of 100 bp
* 101174 103420: contig of 2147 bp in length
* 101174 103420: gap of 100 bp
* 103421 109720: contig of 6200 bp in length
* 103421 109720: gap of 100 bp
* 109721 114256: contig of 4436 bp in length
* 109721 114256: gap of 100 bp
* 114257 119627: contig of 5271 bp in length
* 114257 119627: gap of 100 bp
* 119628 128350: contig of 8623 bp in length
* 119628 128350: gap of 100 bp
* 128351 140183: contig of 11733 bp in length
* 128351 140183: gap of 100 bp
* 140184 147703: contig of 7420 bp in length
* 140184 147703: gap of 100 bp
* 147704 147803: contig of 10297 bp in length
* 147704 147803: gap of 100 bp
* 147804 158100: gap of 100 bp
* 147804 158100: contig of 3782 bp in length
* 158101 162082: gap of 100 bp
* 158201 162082: gap of 100 bp
* 162083 167412: contig of 5330 bp in length.
* 162083 167412: Location/Qualifiers
1. 167412
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP5-859H16"
/clone="RP5-859H16"
/clone="RP5-859H16"
/clone="RP5-859H16"
1. 12019
/note="assembly_fragment:00380
clone_end:826
vector_side:left"
12120..32657
/note="assembly_fragment:00114
fragment_chain:1"
32758..36105
/note="assembly_fragment:00063
fragment_chain:1"
36206..41069
/note="assembly_fragment:01365
fragment_chain:1"
41170..51754
/note="assembly_fragment:01782
fragment_chain:1"
51755..62397
/note="assembly_fragment:01003
fragment_chain:1"
62498..73595
/note="assembly_fragment:00248
fragment_chain:2"
73696..83277
/note="assembly_fragment:01125
fragment_chain:2"
83378..98845
/note="assembly_fragment:01780
fragment_chain:2"
98946..101173
/note="assembly_fragment:01201
fragment_chain:2"
101274..103420
/note="assembly_fragment:01704
fragment_chain:3"
103521..109720
/note="assembly_fragment:00350
fragment_chain:3"
109821..114256
/note="assembly_fragment:01435
fragment_chain:3"
114357..119627
/note="assembly_fragment:00291
fragment_chain:3"

Qy	1	GGACNAGCCCTGACNAGGCC	20	1	Indels	0	Gaps	0
Db	143124	GGTCNAGCCCTGACNAGGCC	143105					

COMMENT
Cambridgehire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Nov 27, 2002 this sequence version replaced gi:24394955.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center -----
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

Together, differences are found these are compared from overlapping clones. variation annotation may not be found in the clone name. Note that the only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the occasion of the clone being a YAC. The following abbreviations are used to associate primary accessions with the feature table with their source databases:

Query Match	92.0%	Score 18.4	DB 10	Length 192657
Best Local Similarity	95.0%	Pred. No. 68		
Matches 19	Conservative 0	Mismatches	1	Indels 0
Gaps				
1	GGACAGGCCCTACAAGCCA	20		
Db	156517	GGACAGGCCCAACAAGCCA	156536	

RESULT 15
 BX294664/c
 LOCUS
 DEFINITION BX294664 193352 bp DNA linear HTG 24-APR-2003
 ACCESSION
 VERSION BX294664.1
 KEYWORDS BX294664.2 GI:29500936
 SOURCE HTG; HTGS_PHASE1; HTGS_CANCELLED.
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 McWay, K.
 Direct Submission
 Submitted (23-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humphrey@sanger.ac.uk
 Clone requests: clonerequests@-----
 On Apr 2, 2003 this HTG Clone request was cancelled.

```

----- quality of at least 30.
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
Center project name: BM384C22
-----
Summary Statistics
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 190946 bases at least Q45
Consensus quality: 191521 bases at least Q40
Consensus quality: 191867 bases at least Q30
Insert size: 192452; sum-of-ctrls
Insert size: 197663; 4.4% error; agarose-fp
Quality coverage: 8.26x in Q20 bases; sum-of-ctrls
Quality coverage: 8.05x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```

Sun Aug 3 09:03:40 2003

UB-09-834-291-10.rge

Search completed: August 1, 2003, 17:32:49
Job time: 83.1573 secs

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 37461: contig of 37461 bp in length
* 37462 37561: gap of 100 bp
* 37562 45950: contig of 8389 bp in length
* 45951 46050: gap of 100 bp
* 46051 98426: contig of 52376 bp in length
* 98427 98526: gap of 100 bp
* 98527 118027: contig of 19501 bp in length
* 118028 118127: gap of 100 bp
* 118128 126520: contig of 8353 bp in length
* 126521 132811: contig of 6191 bp in length
* 132812 132911: gap of 100 bp
* 132912 172432: contig of 39521 bp in length
* 172433 172532: gap of 100 bp
* 172533 185735: contig of 13203 bp in length
* 185736 185835: gap of 100 bp
* 185836 190693: contig of 4858 bp in length
* 190694 190793: gap of 100 bp
* 190794 193352: contig of 2559 bp in length.

FEATURES

source

1. 193352
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-384C22"
/clone_lib="RPCT-23"

misc_feature

1. 37461
/note="assembly_fragment:02244
fragment_chain:1"

misc_feature

vector_end:SP6
37562..45950
/note="assembly_fragment:00826
fragment_chain:1"

misc_feature

46051..98426
/note="assembly_fragment:00297
fragment_chain:1"

misc_feature

98527..118027
/note="assembly_fragment:01245
fragment_chain:1"

misc_feature

118128..126520
/note="assembly_fragment:02356
fragment_chain:1"

misc_feature

126521..132811
/note="assembly_fragment:01763
fragment_chain:1"

misc_feature

132912..172432
/note="assembly_fragment:01130
fragment_chain:1"

misc_feature

172533..185735
/note="assembly_fragment:01150
fragment_chain:1"

misc_feature

185836..190693
/note="assembly_fragment:00830
fragment_chain:1"

misc_feature

190794..193352
/note="assembly_fragment:00599"
vector_end:T7
vector_side:right"

BASE COUNT

58452 a 42616 c 39284 g 52100 t 900 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 193352;
Best Local Similarity 95.0%; Pred No. 68;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGNACGCTGACACCA 20

DB 36135 GGNACGCTGACACCA 36116

Sun Aug 3 09:03:45 2003

US-09-834-291-10.FBT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds
(without alignments)
7038.748 Million cell updates/sec

US-09-834-291-10

Title: 1 gacacagccctgacacagca 20
Sequence: 1 gacacagccctgacacagca 20Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:

1: em_estba:*

2: em_estlm:*

3: em_estlm:*

4: em_estlm:*

5: em_estlm:*

6: em_estlm:*

7: em_estlm:*

8: em_estlm:*

9: gpc_est1:*

10: gpc_est2:*

11: gpc_est3:*

12: gpc_est3:*

13: gpc_est4:*

14: gpc_est5:*

15: em_estlm:*

16: em_estlm:*

17: em_gss_hum:*

18: em_gss_hum:*

19: em_gss_hum:*

20: em_gss_hum:*

21: em_gss_hum:*

22: em_gss_hum:*

23: em_gss_hum:*

24: em_gss_hum:*

25: em_gss_hum:*

26: em_gss_hum:*

27: em_gss_hum:*

28: gpc_est1:*

29: gpc_est2:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Description

Result No.	Score	Query Match Length	DB ID	Description
1	100.0	793	12	B1763679
2	100.0	962	9	AL540709
3	100.0	1089	13	B0072834
4	90.0	486	14	W45225

5	18	90.0	506	10	BG59530
6	18	90.0	753	13	B142933
7	17.4	87.0	639	28	AZ939002
8	17.4	87.0	684	28	AZ450876
9	17.4	87.0	689	29	BX220214
10	17.4	87.0	701	29	CA605171
11	17.4	87.0	745	29	BX179961
12	17.4	87.0	745	29	BX179961
13	17.4	87.0	745	29	BX179961
14	17.4	87.0	745	29	BX179961
15	17.4	87.0	745	29	BX179961
16	17.4	87.0	745	29	BX179961
17	17.4	87.0	745	29	BX179961
18	17.4	87.0	745	29	BX179961
19	17.4	87.0	745	29	BX179961
20	17.4	87.0	745	29	BX179961
21	17.4	87.0	745	29	BX179961
22	17.4	87.0	745	29	BX179961
23	17.4	87.0	745	29	BX179961
24	17.4	87.0	745	29	BX179961
25	17.4	87.0	745	29	BX179961
26	17.4	87.0	745	29	BX179961
27	17.4	87.0	745	29	BX179961
28	17.4	87.0	745	29	BX179961
29	17.4	87.0	745	29	BX179961
30	17.4	87.0	745	29	BX179961
31	17.4	87.0	745	29	BX179961
32	17.4	87.0	745	29	BX179961
33	17.4	87.0	745	29	BX179961
34	17.4	87.0	745	29	BX179961
35	17.4	87.0	745	29	BX179961
36	17.4	87.0	745	29	BX179961
37	17.4	87.0	745	29	BX179961
38	17.4	87.0	745	29	BX179961
39	17.4	87.0	745	29	BX179961
40	17.4	87.0	745	29	BX179961
41	17.4	87.0	745	29	BX179961
42	17.4	87.0	745	29	BX179961
43	17.4	87.0	745	29	BX179961
44	17.4	87.0	745	29	BX179961
45	17.4	87.0	745	29	BX179961

ALIGNMENTS

RESULT 1
LOCUS B1763679/c
DEFINITION B1763679
ACCESSION B1763679
VERSION B1763679.1
KEYWORDS B1763679.1
SOURCE B1763679.1
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE NIH-MGC
JOURNAL NIH-MGC
COMMENT NIH-MGC

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rstraus@nihs.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLIML at:
http://image.llnl.gov
Plate: LIML1474
High quality sequence stop: 786.

FEATURES

Location/Qualifiers

1..962

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5189752"

/lab_host="DH10B"

/clone_lib="NIH MGC 116"

/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; site: 1: NotI; site 2: EcoRV (destroyed); RNA source anonymous pool of 3 clones, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC library."

BASE COUNT 163 a 265 c 208 g 137 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 12; Length 793;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAAAGCCCTGACAAAGCA 20
398 GGACAAAGCCCTGACAAAGCA 379

RESULT 2

AL540709/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AL540709 962 bp mRNA linear EST 12-MAY-2003
5-PRIME, mRNA sequence.
AL540709
EST.
GI:30544172
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 962)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12871113.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5554.r for
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE002YXN18
Feng Liang Email: fliang@life.techn.com URL: <http://fulllength.invitrogen.com/>
Faraday Avenue Genoscope sequence ID: CS0DE002YXN18
Location/Qualifiers
1..962

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE002YXN18"
/issue_type="PLACENTA"
/note="Vector: PCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
the Not I and EcoRV sites of the PCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT

211 a 302 c 264 g 161 t 4 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 9; Length 962;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAAAGCCCTGACAAAGCA 20
94 GGACAAAGCCCTGACAAAGCA 75

RESULT 3

B0072834

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

B0072834 1089 bp mRNA linear EST 02-APR-2002
5', mRNA sequence.
B0072834
EST.
GI:19901880
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1089)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: Life Technologies, Inc.
cDNA library Preparation: Life Technologies, Inc.
DNA sequencing by: Agencourt Bioscience Corporation (LNL)
found through the I.M.A.G.E. Consortium
<http://image.llnl.gov>
Plate: LLM12796 row: 1 column: 21
High quality sequence stop: 645.

FEATURES

source

Location/Qualifiers
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5756324"
/issue_type="Leukocyte"
/lab_host="DH10B"
/clone_lib="NIH MGC 118"
/note="Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH MGC library."

BASE COUNT

280 a 263 c 333 g 212 t 1 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 13; Length 1089;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAAAGCCCTGACAAAGCA 20
369 GGACAAAGCCCTGACAAAGCA 388

RESULT 4

W45225

LOCUS

DEFINITION

W45225

W45225 486 bp mRNA linear EST 10-OCT-1996
zc23c09.r1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
Clone IMAGE:123152.5									
PHOC (HUMAN); mRNA sequence.									
W45225									
W45225.1	GI:1329306								
EST.									
Homo sapiens (human)									
Homo sapiens									
Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.									
1 (bases 1 to 486)									
Haller, L., Clark, N., Dubucq, T., Elliston, K., Hawkins, M., Holman									
, M., Hultman, M., Kueba, T., Le, M., Lennon, G., Morris, M., Parsons, J.,									
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston									
, R., Williamson, A., Wohlmann, P. and Wilson, R.									
The Washu-Merck EST Project									
Unpublished									
Contact: Wilison RK									
Washington University School of Medicine									
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108									
Tel: 314 286 1800									
Fax: 314 286 1810									
Email: est@wuston.wustl.edu									
This clone is available royalty-free through LNL; contact the									
IMAGE Consortium (info@image.lnl.gov) for further information.									
Insert Length: 972 Std Error: 0.00									
Seq primer: mob. RBGA+RT									
High quality sequence stop: 296.									
Location/Qualifiers									
1..486									
/organism="Homo sapiens"									
/mol_type="mRNA"									
/db_xref="GDB:1254664"									
/db_xref="taxon:9606"									
/clone="IMAGE:323152"									
/tissue_type="senescent fibroblast"									
/lab_host="DH10B (ampicillin resistant)"									
/clone_1lb="Soares_senescent fibroblasts NBHSF"									
/note="Vector: pVT3D (Pharmacia) with a modified									
polylinker V type: phagemid; Site 1: Not I; Site 2: Eco RI									
, 1st strand cDNA was primed with a Not I - oligo(dT)									
primer [5',									
TGTTACCACTCTGAGTGGAGCGCGCAATTTTTTTTTTTT 3']									
double-stranded cDNA was size selected, ligated to Eco RI									
adapters (Pharmacia), digested with Not I and cloned into									
the Not I and Eco RI sites of a modified pVT3D vector									
(Pharmacia). Library went through one round of									
normalization to a Cot = 5. Library constructed by Bento									
Soares and M. Fatima Bonaldo."									
2 others									
BASE COUNT									
120 a 129 c 140 g 95 t									
ORIGIN									
Query Match									
Best Local Similarity 100.0%; Pred. No. 3.4e-02;									
Matches 18; Conservative 0; Mismatches 0; Gaps 0;									
2 GACAGCCCTGACAGCC 19									
265 GACAGCCCTGACAGCC 282									
RESULT 5									

REFERENCE	1 (bases 1 to 506)
AUTHORS	Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Afjoka, J.A., White, M., Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kuchaba, T., Thelaising, B., Bowers, Y., Gibbons, M., Rietter, E., Bennett, J., Franklin, C., Tsagaratsehlvili, R., Ronko, I., Kennedy, S., Maguire, L., Waterston, R., and Wilson, R.
TITLE	Toxoplasma EST Project
JOURNAL	Unpublished
COMMENT	Contact: Clifton, S. Toxoplasma EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: toxo@wustl.wustl.edu Contact David Sibley (toxoes@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability Seq primer: -40RP from Gibco High quality sequence stop: 371. location/Qualifiers 1. 506 /organism="Toxoplasma gondii"
FEATURES	
SOURCE	

ORIGIN 86 a 151 C 170 S
 BASE COUNT
 Query Match 90.0%; Score 18; DB 10; Length 506;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0;
 Oy 1 GGACACAGCCCTGCACACG 18
 Db 18 GGACACAGCCCTGCACACG 1
 RESULT 6
 BU142933/c 753 bp mRNA linear EST 25-NOV-2002
 LOCUS 603135537F1 CSEQCH25 Gallus gallus clone CHEST11919 5', mRNA
 DEFINITION
 sequence.
 BU142933
 BU142933.1 GI:25358894
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Gallus gallus (chicken)
 Gallus gallus
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Accipitriformes; Phasianidae;
 Phasianidae; Gallus.
 1 (bases 1 to 753)
 Boardman, P.E., Sanz-Exquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Ford, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken CDNAS
 Curr. Biol. 12 (22), 1965-1969 (2002)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 2235534
 12445392


```

Class: plasmid ends
High quality sequence stop: 619
Location/Qualifiers
    source
FEATURES
    1. .619

```

BASE COUNT	FEATURES	source
198 a	Class: plasmid ends High quality sequence stop: 619.	
116 c	Location/Qualifiers	
87 g	1..619	
218 c	/organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="TUGC2M0197J18" /sex="Female" /lab_host="E. coli strain XL10-Gold, T1-resistant, F-" /clone_idb="Mouse.10kb plasmid TUGC2M library" /note="Vector: PMD42NV, Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211[gb AF129072.1]), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and adapted vector DNA, sheared, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."	

BASE COUNT	ORIGIN	Score	DB	Length
198 a	116 c	87.0%	17.4	619
		94.7%	Pred. No. 7, 1e+02	
Matches	18; Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;
QY	1 GGACAGCCTGACAGCC	19		
Db	595 GGACAGCCTGACAGCC	577		

RESULT 8	AZ450876	684 bp	DNA	linear	SSS 04-OCT-2000
LOCUS	1M0243M1.R Mouse 10kb plasmid UGCGM library Mus musculus genomic				
DEFINITION	clone UGCGM0449M17 R, genomic survey sequence.				
ACCESSION	AZ450876				
VERSION	AZ450876.1	GI:10606116			
KEYWORDS	SSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.				
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb				
JOURNAL	plasmid inserts				
COMMENT	Unpublished				
CONTACT	Robert D. Dunn				

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0248 row: M column: 17

FEATURES

SOURCE

Seq primer: CACACAGAAAACAGTATGACC
 Class: plasmid ends
 High quality sequence stop: 684.
 Location/Qualifiers

1..684
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0249M17"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb Plasmid UUCG1M library"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/nanres/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|9b|AF12072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, and transformed into chemically-competent E.coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

180 a 202 c 160 g 142 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 684;
 Best Local Similarity 94.7%; Pred. No. 7, 3e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCA 20
 Db 611 GACAGCACTGACAGCA 629

RESULT 9 BX220214/c 689 bp DNA linear GSS 29-JAN-2003
 LOCUS BX220214
 DEFINITION Danio rerio genomic clone DKEY-269M11, genomic survey sequence.
 ACCESSION BX220214
 VERSION BX220214.1 GI:28052100
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 689)
 Humphray,S.J., Huckle,E. and Durham,J.L.
 Direct Submission
 Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humphray@sanger.ac.uk unpublished
 This sequence was generated from the SP6 end of BAC 269M11. 269M11
 is part of the Daniokey BAC library created by R. Plaetzer and N.V.
 Keygene. Further details:
 http://www.sanger.ac.uk/projects/D_rerio/
 Location/Qualifiers

FEATURES

SOURCE

1..689
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-269M11"
 /tissue_type="Testis"

BASE COUNT

/note="vector pindigBAC-536"
 180 a 144 c 144 g 221 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 29; Length 689;
 Best Local Similarity 94.7%; Pred. No. 7, 4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCA 20
 Db 419 GACAGCCCTGACAGCA 401

RESULT 10

CA605171/c 701 bp mRNA linear EST 21-NOV-2002
 LOCUS CA605171
 DEFINITION w1.pK0046.e7 w1 Triticum aestivum cDNA clone w1.pK0046.e7 5' end
 , mRNA sequence.
 ACCESSION CA605171
 VERSION CA605171.1 GI:25160333
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Triticum.
 1 (bases 1 to 701)
 Tingey,S.V., Powell,M., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
 Miao,G., Caraher,N. and Hanafey,M.K.
 Dupont Wheat cDNA Sequence
 Unpublished
 Contact: Scott V. Tingey
 Crop Genetics
 E. I. Dupont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@USA.dupont.com
 Seq primer: M3.
 Location/Qualifiers

FEATURES

SOURCE

1..701
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="w1.pK0046.e7"
 /tissue_type="root"
 /clone_lib="w1"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
 XhoI; Wheat (Triticum aestivum L.) root; 7 day old
 seedling, light grown"

BASE COUNT

199 a 188 c 131 g 156 t 27 others

ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 701;
 Best Local Similarity 94.7%; Pred. No. 7, 4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACAGCCCTGACAGCC 19
 Db 255 GACAGCCCTGACAGCC 237

RESULT 11

BX179961/c 745 bp DNA linear GSS 28-JAN-2003
 LOCUS BX179961
 DEFINITION Danio rerio genomic clone DKEY-183M20, genomic survey sequence.
 ACCESSION BX179961
 VERSION BX179961.1 GI:28011764
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

FEATURES

SOURCE

1..745
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-183M20"
 /tissue_type="Testis"

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 745)
Humphrey, S.J., Huckle, E. and Durham, J.L.
Direct Submission
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphrey@sanger.ac.uk Unpublished
This sequence was generated from the SPE end of BAC 183W20. 183W20
is part of the Daniokey BAC library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_danio/
Location/Qualifiers
1..745

BASE COUNT
ORIGIN

209 a 150 c 151 g 235 t

Query Match 87.0%; Score 17.4; DB 29; Length 745;
Best Local Similarity 94.7%; Pred. No. 7.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCCA 20
|||||
419 GACAGCCCTGACAGCCA 401

RESULT 12
LOCUS

T62311

DEFINITION T62311 331 bp mRNA linear EST 27-AUG-1998
EST0074 TGRH Tachyzoite cDNA Toxoplasma gondii clone CGD008
5' end, mRNA sequence.

ACCESSION T62311
VERSION T62311.1 GI:665753
KEYWORDS EST.
SOURCE Toxoplasma gondii
ORGANISM Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.

REFERENCE 1 (bases 1 to 331)
AUTHORS Wan, K.-L., Blackwell, J.M. and Ajioke, J.W.
TITLE Toxoplasma gondii expressed sequence tags: insight into tachyzoite
gene expression
Mol. Biochem. Parasitol. 75 (2), 179-186 (1995)
8992316

JOURNAL MEDLINE
PUBMED
COMMENT Contact: Ajioke JW
Laboratory for Parasite Genome Analysis
Cambridge University
Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK.
Tel: 01223333923
Fax: 01223333923
Email: jajioke@gmp.mrc.ac.uk
Clone information and library can be obtained from Jim Ajioke
jw@mol.bio.cam.ac.uk
Seq primer: SK.

FEATURES
source

Location/Qualifiers

1..331
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="RH"
/db_xref="taxon:5811"
/clone="cgd008"
/lab_host="XLI-Blue MRF"
/clone_lib="TGRH Tachyzoite cDNA"
/note="Vector: Lambda ZAP; Site 1: EcoRI; Site 2: XhoI;
Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
constructed by K.L. Wan, Cambridge University. CDNA's were

BASE COUNT
ORIGIN

synthesized from polyA RNAs by oligo d(T) priming and
directionally cloned into the pCR2.1 to XhoI sites of the
Lambda ZapII vector using the ZAP-cDNA synthesis kit
(Stratagene). WARNING: the library contains a small
percentage of cDNAs derived from the human host cells."
70 a 94 c 80 g 87 t

Query Match 85.0%; Score 17; DB 14; Length 331;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAG 17
|||||
314 GGACAGCCCTGACAG 330

RESULT 13
LOCUS

A2781046

DEFINITION A2781046 429 bp DNA linear GSS 16-FEB-2001
2M0018K23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0018K23 R, genomic survey sequence.

ACCESSION A2781046
VERSION A2781046.1 GI:12913345
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 429)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausen, A.
and Wright, D.J. Weiss, R. Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
COMMENT

Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 306, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5605
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0018 row: K column: 23
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 429.

FEATURES
source

Location/Qualifiers

1..429
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0018K23"
/sex="Male"
/lab_host="E. Coli strain XLI-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: pMD229v. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (914732114|9b|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 125 a 88 c 126 g 90 t

Query Match 85.0%; Score 17; DB 28; Length 429;

Best Local Similarity 100.0%; Pred. No. 9.6e+02; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAG 17

Db 227 GGACAGCCTGACAG 243

RESULT 14 AA195416 531 bp mRNA linear EST 06-AUG-1997

LOCUS z36f10.s1 Soares_NbHMPu_S1 Homo sapiens cDNA clone IMAGE:665515

DEFINITION 3', mRNA sequence.

ACCESSION AA195416 GI:1785109

VERSION AA195416.1 GI:1785109

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 531) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kucala, T., Lacy, M., Le, N., Lennon, G., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson R.

Washu-Merck EST Project 1997

Unpublished

CONTACT: Wilson R. Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800 Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LML; contact the

IMAGE Consortium (info@image.lml.gov) for further information.

Insert Length: 1020 Std Error: 0.00

High quality sequence stop: 436.

Location/Qualifiers

1..531

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:5427547"

/db_xref="taxon:9606"

/clone="IMAGE:665515"

/tissue_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab_host="DH10B"

/note="Organ: mixed (see below); Vector: pTRT3D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NBH4, pregnant uterus

NbHPU, and fetal heart NbH19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools

consisted of 1 M.A.G.E. clones 260232-265223,

340488-345479, and 484488-489479."

BASE COUNT 111 a 110 c 167 g 141 t 2 others

Query Match 84.0%; Score 16.8; DB 9; Length 531;

Best Local Similarity 90.0%; Pred. No. 1.3e+03; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAGCCA 20

Db 158 GGACAGCCTGACAGCCA 139

RESULT 15

LOCUS AM11236 592 bp mRNA linear EST 27-OCT-1999

DEFINITION x160c04.x1 NCI_CGAP Gas4 Homo sapiens cDNA clone IMAGE:2622438 3',

similar to SW:TF1B HUMAN Q13263 TRANSCRIPTION INTERMEDIARY FACTOR

1-BETA, mRNA sequence.

ACCESSION AM11236 GI:6132843

VERSION AM11236.1 GI:6132843

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 592) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

Tumor Gene Index

Unpublished

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmett-Buck, M.D., Ph.D.

CDNA library preparation: Life Technologies, Inc.

DNA sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LML at:

www.bio.lml.gov/bdrip/image/image.html

Seq primer: -40UP from G1bco

High quality sequence stop: 423.

Location/Qualifiers

1..592

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2622438"

/tissue_type="poorly differentiated adenocarcinoma with

signet ring cell features"

/lab_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.69 kb. Life Technologies catalog #:

11549-011"

BASE COUNT 133 a 146 c 194 g 115 t 4 others

Query Match 84.0%; Score 16.8; DB 9; Length 592;

Best Local Similarity 90.0%; Pred. No. 1.3e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAGCCA 20

Db 590 GGACAGCCTGACAGCCA 571

Search completed: August 1, 2003, 20:42:58

Job time: 72.4754 secs

XX WPI; 2000-162245/15.

XX Novel receptor DNA useful for identifying apoptosis-modulating
XX substances potentially useful for cancer chemotherapy -
XX Claim 2; Fig 4; 12pp; German.

CC This invention describes a novel p53-binding region of a human CD95
CC receptor DNA molecule. The p53-binding region, or a vector containing
CC it, can be used to screen for apoptosis-modulating substances
CC potentially useful for cancer chemotherapy. This sequence represents a
CC fragment of the human CD95 receptor intron 1 which contains a p53 binding
CC region described in the method of the invention.

SO Sequence 266 BP; 49 A; 72 C; 110 G; 35 T; 0 other;

Query Match Best Local Similarity 100.0%; Score 20; DB 21; Length 266;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACACAGCCCTGACACGCCA 20
DB 160 GACACAGCCCTGACACGCCA 179

RESULT 2
ABX91757/c ID ABX91757 standard; cDNA; 516 BP.

AC ABX91757;

DT 07-MAY-2003 (first entry)

DE Murine gene trapped sequence (GTS) SEQ ID NO 1101.

KW Murine; mouse; gene trap technology; gene trapped sequence; GTS;
KW gene identification; functional genomic analysis; gene discovery;
KW gene expression analysis; cross species hybridisation analysis;
KW antisense inhibition; gene targeting; gene; ss.

OS Mus sp.

PN US2002161207-A1.

PD 31-OCT-2002.

PF 30-NOV-2000; 2000US-0728444.

PR 01-DEC-1999; 99US-168360P.

PA (FRIE/) FRIEDRICH G.
PA (ZAMB/) ZAMBROWICZ B.
PA (SAND/) SANDS A T.

PI Friedrich G, Zambrowicz B, Sands AT;

DR WPI; 2003-288124/28.

PT New murine polynucleotides comprising gene trapped sequences, useful in
PT functional genomic analysis, in the development of new therapeutic or
PT diagnostic agents, for diagnostic gene expression analysis or for
PT genetic manipulations -
XX Claim 2; SEQ ID NO 1101; 29pp; English.

CC The present invention relates to novel murine cDNAs produced using
CC gene trap technology. The OMNIBANK gene trapped sequences (GTSs)
CC are individually identified novel genes, and are useful in functional
CC genomic analysis, in the discovery and development of new therapeutic
CC and diagnostic agents, for gene discovery, for diagnostic gene
CC expression analysis, for cross species hybridisation analysis, and for
CC genetic manipulations such as antisense inhibition or gene targeting.

CC The polynucleotides of the invention are also useful for isolating
CC cDNAs, genomic clones or full-length genes/polynucleotides, or their
CC homologues, heterologues, paralogues or orthologues, that are capable
CC of hybridising to one or more of the new murine polynucleotide
CC sequences. The polynucleotides are also useful for identifying the
CC coding regions of the murine genome, and as hybridisation probes.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/psipdsidentry.html.

SO Sequence 516 BP; 163 A; 90 C; 118 G; 140 T; 5 other;

Query Match Best Local Similarity 82.0%; Score 16.4; DB 25; Length 516;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACACAGCCCTGACACGCC 19
DB 100 GACACAGCCCTGACACGCC 83

RESULT 3
ABZ74553 ID ABZ74553 standard; DNA; 8243 BP.

AC ABZ74553;

DT 12-MAY-2003 (first entry)

DE Secreted protein gene 346 genomic fragment HTOH021, SEQ ID NO:1700.

KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytotoxic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antineoplastic; vulnery; chromosome 15q24-25; gene; ds.

OS Homo sapiens.

PN WO200277013-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002WO-US093370.

PR 27-MAR-2001; 2001US-278650P.

PR 12-SEP-2001; 2001US-0950082.

PR 12-SEP-2001; 2001US-0950083.

PA (HDMA-) HDMA GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2003-040578/03.

PT New human secreted proteins and nucleic acids, useful for detecting or
PT treating cancer or other hyperproliferative disorders, autoimmune
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -
XX Disclosure; Page 2282-2284; 2474pp; English.

CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cellular differentiation, cell
CC migration, proinflammatory activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody

fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein genomic fragment referred to in the disclosure of the invention.

Sequence 8243 BP; 2717 A; 1618 C; 1711 G; 2197 T; 0 other;

Query Match 82.0%; Score 16.4; DB 25; Length 8243;
Best Local Similarity 94.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCC 19
DB 8148 GACAGCCCTGACAGCC 8165

RESULT 4
AAK81012/C
ID AAK81012 standard; DNA; 12710 BP.

AC AAK81012;
DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35824.
XX Human immune/haematopoietic; immune/haematopoietic antigen; cancer;
KM cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX MO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216680.

XX 07-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234998.

PR 25-SEP-2000; 2000US-0235484.

PR 26-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236370.

PR 29-SEP-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241825.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246509.

08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0251719.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 DR
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PS
 XX
 PS Disclosure; SEQ ID NO 35824; 3071bp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX Sequence 12710 BP; 3496 A; 2567 G; 4072 T; 0 other;

Query Match 82.0%; Score 16.4; DB 22; Length 12710;
 Best Local Similarity 94.4%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 2 GACAAGCCCTGAGCAAGCC 19

Db 96 GACAAGCCCTGAGCAAGCC 79

RESULT 5
 AB274552
 ID AB274552 standard; DNA; 12710 BP.
 XX
 AC AB274552;
 XX
 DT 12-MAY-2003 (first entry)
 XX
 DE Secreted protein gene 346 genomic fragment HMOH021, SEQ ID NO:1699.

Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antinaemic; vulnery; chromosome 15q24-25; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277013-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-US093370.
 XX
 PR 27-MAR-2001; 2001US-278650P.
 PR 12-SEP-2001; 2001US-0950082.
 PR 12-SEP-2001; 2001US-0950083.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX WPI; 2003-040578/03.
 DR
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -
 XX
 PS Disclosure; Page 2278-2281; 2474bp; English.

AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
 CC AB273698-AB274687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins.
 CC The use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC associated proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prohormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 XX
 XX Sequence 12710 BP; 4072 A; 2575 G; 3496 T; 0 other;

Query Match 82.0%; Score 16.4; DB 25; Length 12710;
 Best Local Similarity 94.4%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCTGACAGCC 19
 |||||
 DB 12615 GACAGCCTGACAGCC 12632

RESULT 6
 ABX52814/C
 ID ABX52814 standard; cDNA; 433 BP.

XX ABX52814;

DT 25-FEB-2003 (first entry)

XX Bovine EST associated with lactation/muscle/fat deposition #2743.

DE Bovine; EST; expressed sequence tag; lactation; LMPD;

KW muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137160-A1.

XX 26-SEP-2002.

PD 26-OCT-2001; 2001US-0983965.

PF 17-DEC-1998; 98US-113678P.

PR 15-DEC-1999; 99US-0465231.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

PI Byatt JC, Mathialagan N, Tao N, Warren WC,

XX WPI; 2003-102386/09.

XX Purified nucleic acid molecules, useful for genome mapping, gene

PT identification and analysis, cattle breeding or preparation of

PT constructs for cattle gene expression and genetically improved cattle -

XX Claim 2; SEQ ID NO 2743; 38pp; English.

PS The invention relates to a purified nucleic acid molecule associated with

XX lactation or muscle and fat deposition (designated LMPD), derived to a

CC from cattle, and the LMPD nucleic acid can specifically hybridise to a

CC second nucleic acid molecule comprising any of 5912 nucleotide

CC sequences, appearing as ABX50072-ABX55983, or complements of them.

CC Also included are: (1) a transformed cell having a nucleic acid

CC comprising an LMPD nucleic acid linked to a promoter and a 3' end

CC transcription sequence that functions in the cell to cause termination of a

CC molecule in a bovine cell or tissue, where hybridisation between the

CC nucleic acid (comprising any of the 5912 nucleic acid sequences or its

CC complement or fragment) with a complementary nucleic acid molecule

CC obtained from the bovine cell or tissue, where hybridisation between the

CC marker nucleic acid and the complementary nucleic acid permits the

CC detection of the molecule; and (b) detecting the level or pattern of the

CC complementary nucleic acid, where the detection of the complementary

CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMPD nucleic acid is used for determining a level or pattern

CC of a molecule in a bovine cell or tissue. It is useful for genome

CC mapping, gene identification and analysis, cattle breeding, preparation

CC of constructs for use in cattle gene expression, or for genetically

CC improving cattle. The present tag nucleic acids.

CC LMPD EST (expressed sequence tag) nucleic acids.

CC Note: The present sequence was not shown in the USPTO web site:

CC was obtained in electronic format from the USPTO web site:

XX Sequence 433 BP; 86 A; 88 C; 159 G; 100 T; 0 other;
 SQ Query Match 80.0%; Score 16; DB 25; Length 433;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0

QY 4 CAAGCCTGACAGCC 19
 |||||
 DB 102 CAAGCCTGACAGCC 87

RESULT 7
 ABX27459
 ID ABX27459 standard; cDNA; 165 BP.

XX ABX27459;

DT 11-FEB-2003 (first entry)

DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #9516.

KW Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; se; inflammation;

KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;

KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;

KW complex carbohydrate; gene replacement therapy; immunosuppressive;

KW anti-inflammatory; antiarthritic; antibacterial; cerebroprotective;

KW antiasthmatic; vasotropic.

XX Homo sapiens.

XX US2002110548-A1.

XX 15-AUG-2002.

PD 11-JUN-2001; 2001US-0878574.

PF 22-NOV-1996; 96US-0753233.

PR 03-DEC-1997; 97US-0984246.

PR 09-SEP-1998; 98US-0149674.

PR 14-JUN-1999; 99US-0333177.

XX (GENVY) GENETICS INST INC.

PI Sullivan F, Kriz R, Kumar R;

XX WPI; 2003-066673/06.

XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)

PT peptide, for manufacturing complex carbohydrates, or as targets for

PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant

PT rejection

XX Disclosure; SEQ ID NO 9518; 6pp; English.

PS The invention relates to a composition comprising a human GDP-mannose

XX 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying

CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation

CC in a mammalian subject and for treating or ameliorating diseases affected

CC by the level of cellular fucosylation or diseases include arthritis,

CC fucosylation of glycoconjugates. These diseases include stroke or

CC transplant rejection, asthma, or a polynucleotide encoding it is also

CC useful for manufacturing complex carbohydrates and as targets for

CC screening small molecule antagonists of the activity of the enzyme. The

CC polynucleotide is useful in developing an assay for defects in the

CC enzyme, as well as in gene replacement therapy. Sequences

CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding

CC human GM4,6D peptides of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC at seqdata.uspto.gov/sequence.html.

SO Sequence 165 BP; 34 A; 66 C; 42 G; 23 T; 0 other;

Query Match 79.0%; Score 15.8; DB 25; Length 165;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCACAGCCCTGCACAGCC 19
Db 131 GCACAGCCCTGCACAGCC 149

RESULT 8

AB241531
ID AB241531 standard; DNA; 261 BP.

AC AB241531;

DT 07-MAR-2003 (first entry)

DE N. gonorrhoeae nucleotide sequence SEQ ID 7651.

KW Antibacterial; infection; vaccine; gene therapy; gene; ds.

OS Neisseria gonorrhoeae.

PN MO200279243-A2.

PD 10-OCT-2002.

PR 12-FEB-2002; 2002WO-IB02069.

PR 12-FEB-2001; 2001GB-0003424.

PA (CHIR-) CHIRON SPA.

PI Fontana MR, Pizsa M, Massignani V, Monaci E;

DR WPI; 2003-058415/05.

DR F-PSDB; ABP80561.

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
medicament for treating or preventing N. gonorrhoeae infection -
PS Disclosure; Page 746; 815pp; English.

CC The present invention relates to proteins from Neisseria gonorrhoeae.
CC Antibodies that specifically bind to the proteins encoding the proteins and
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records AB237706-AB242016 represent nucleic acid
CC molecules of the invention.

SO Sequence 261 BP; 53 A; 71 C; 65 G; 72 T; 0 other;

Query Match 79.0%; Score 15.8; DB 25; Length 261;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCACAGCCCTGCACAGCCA 20
Db 153 GCACAGCCCTGCACAGCCA 171

RESULT 9

AA211974/c

ID AA211974 standard; DNA; 273 BP.

AC AA211974;

DT 08-OCT-1999 (first entry)

DE Neisseria meningitidis strain A complete ORF37 sequence.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.

OS Neisseria meningitidis.

PN MO9924578-A2.

PD 20-MAR-1999.

PR 09-OCT-1998; 98WO-IB01665.

PR 01-SEP-1998; 98GB-0019016.

PR 06-NOV-1997; 97GB-0023516.

PR 14-NOV-1997; 97GB-0024190.

PR 18-NOV-1997; 97GB-0024386.

PR 27-NOV-1997; 97GB-0025158.

PR 10-DEC-1997; 97GB-0026147.

PR 14-JAN-1998; 98GB-000759.

PA (CHIR-) CHIRON SPA.

PI Grandi G, Massignani V, Pizsa M, Rappuoli R, Scarlato V;

DR WPI; 1999-327407/27.

DR P-PSDB; AAY8501.

PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
diagnosis, treatment and prevention of infection
PS Claim 3; Page 61; 524pp; English.

CC Nucleotide sequences AA211972-212358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.

SO Sequence 273 BP; 75 A; 67 C; 83 G; 48 T; 0 other;

Query Match 79.0%; Score 15.8; DB 20; Length 273;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCACAGCCCTGCACAGCCA 20
Db 217 GCACAGCCCTGCACAGCCA 199

RESULT 10

AA253178

ID AA253178 standard; DNA; 288 BP.

AC AA253178;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 091 partial DNA sequence SEQ ID NO:307.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy; ds.

OS Neisseria meningitidis.

PN MO9957280-A2.

PD 11-NOV-1999.

PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000755.

PA (CHIR-) CHIRON SPA.

XX Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX WPI, 1999-327407/27.
DR P-PSDB; AAY38502.

XX Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
XX diagnosis, treatment and prevention of infection
PS Claim 3; Page 61; 524pp; English.

CC Nucleotide sequences AAZ11972-212358 represent open reading frames
CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode
CC antigenic proteins (see AAY38499-X38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis, their
CC prevention (as vaccines) or treatment of *Neisseria* infections,
CC such as meningitis, septicemia and gonorrhea. Both organisms,
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.

XX Sequence 381 BP; 106 A; 84 C; 114 G; 77 T; 0 other;

Query Match 79.0%; Score 15.8; DB 20; Length 381;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACAGCGCTGACAGCCCA 20
Db 325 GACAGCGCTGACAGCCCA 307

RESULT 13

AAKS8992
ID AAKS8992 standard; cDNA, 409 BP.

XX AAKS8992;

DT 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:4052.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytosolic; gene therapy; vaccine; metastasis; ss.

XX Homo sapiens.

PN W0200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0188874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 07-JUN-2000; 2000US-0205515.

PR 28-JUN-2000; 2000US-0209467.

PR 30-JUN-2000; 2000US-0214886.

PR 07-JUL-2000; 2000US-0215135.

PR 11-JUL-2000; 2000US-0216647.

PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225256.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226586.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231245.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 20-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.

Qy	1	2	Indels	Gaps	0
Db	GGACAGCCCTTGACAAAGCC 19				
	558 GGACATGCCCGGACAAAGCC 576				

ID AAK74243 standard; DNA; 749 BP.
XX AAK74243;
AC
XX
DT 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29055.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
XX MO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US01354.
XX
PF 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190706.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0218290.
PR 14-JUL-2000; 2000US-0218296.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225770.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 18-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 22-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-024617.
PR 01-NOV-2000; 2000US-0246176.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

```

PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251855.
PR 08-DEC-2000; 2000US-0251855.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX PS Disclosure; SEQ ID NO 29055; 3071pp + Sequence listing; English.
XX
XX CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent the
XX diagnosis and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 749 BP; 201 A; 185 C; 137 G; 226 T; 0 other;
XX
XX Query Match 79.0%; Score 15.8; DB 22; Length 749;
XX Best Local Similarity 89.5%; Pred. No. 2.7e+02;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 GGACACAGCCCTGACAGCC 19
XX |||||
XX Db 558 GGACATGCCCGACACAGCC 576

```

Search completed: August 1, 2003, 13:55:14
 Job time : 7.69868 secs

Sun Aug 3 09:03:43 2003

us-09-834-291-10.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 : Search time 5.18607 Seconds
(without alignments)
7955.924 Million cell updates/sec

US-09-834-291-10

Title: 20
Perfect score: 1 ggcacagcctgacacgcca 20
Sequence: 1 ggcacagcctgacacgcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq: *
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq: *
15: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *
16: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	US-09-834-291-5	Sequence 5, Appli
2	20	100.0	20	US-09-834-291-10	Sequence 10, Appli
3	20	100.0	20	US-09-834-291-24	Sequence 24, Appli
4	20	100.0	266	US-09-834-291-32	Sequence 32, Appli
5	20	100.0	720	US-09-834-291-3	Sequence 3, Appli
6	20	100.0	2380	US-09-834-291-4	Sequence 4, Appli
7	20	100.0	2827	US-09-834-291-1	Sequence 1, Appli
8	20	100.0	3212	US-09-834-291-6	Sequence 6, Appli
9	18.4	92.0	20	US-09-834-291-7	Sequence 7, Appli
10	16.8	84.0	20	US-09-834-291-9	Sequence 9, Appli
11	16.8	84.0	20	US-10-027-632-174763	Sequence 174763,
12	16.8	84.0	2940917	US-10-027-632-174763	Sequence 1101, Ap
13	16.4	82.0	516	US-09-728-444-1101	Sequence 2743, Ap
14	16	80.0	433	US-09-983-965-2743	Sequence 9518, Ap
15	15.8	79.0	165	US-09-878-574-9518	Sequence 87503, A
16	15.8	79.0	487	US-10-027-632-87503	

17	15.8	79.0	487	13	US-10-027-632-316681	Sequence 316681,
18	15.8	79.0	88421	10	US-09-976-059-1	Sequence 1, Appli
19	15.8	79.0	155074	10	US-10-026-188-6	Sequence 6, Appli
20	15.8	79.0	3309400	13	US-09-738-762-1	Sequence 1, Appli
21	15.8	79.0	9025608	14	US-10-156-761-1	Sequence 3975, Ap
22	15.4	77.0	272	10	US-09-983-965-3975	Sequence 493, Ap
23	15.4	77.0	401	9	US-09-795-668-493	Sequence 493, App
24	15.4	77.0	401	9	US-09-946-807-493	Sequence 966, App
25	15.4	77.0	401	10	US-10-027-632-263366	Sequence 567, App
26	15.4	77.0	634	13	US-10-027-632-966	Sequence 3925, Ap
27	15.4	77.0	673	10	US-09-823-830A-557	Sequence 3926, Ap
28	15.4	77.0	845	10	US-09-764-877-3925	Sequence 127733,
29	15.4	77.0	860	10	US-09-764-877-3925	Sequence 127733,
30	15.4	77.0	872	13	US-10-027-632-127733	Sequence 127733,
31	15.4	77.0	872	13	US-10-027-632-127733	Sequence 127733,
32	15.4	77.0	2155	13	US-10-027-632-102861	Sequence 102861,
33	15.4	77.0	2155	13	US-10-027-632-102862	Sequence 258, App
34	15.4	77.0	2155	14	US-10-128-714-258	Sequence 528, App
35	15.4	77.0	3604	14	US-10-900-449A-3	Sequence 1, Appli
36	15.4	77.0	34668	11	US-09-795-668-1	Sequence 1, Appli
37	15.4	77.0	1503841	9	US-09-795-668-1	Sequence 1862, Ap
38	15.4	77.0	1503841	10	US-09-946-807-1	Sequence 2134, Ap
39	15.4	77.0	1503841	10	US-09-983-965-1862	Sequence 2161, Ap
40	15.4	77.0	424	10	US-09-983-965-2134	Sequence 187538,
41	15.2	76.0	428	10	US-09-783-590-2161	Sequence 187539,
42	15.2	76.0	502	10	US-09-783-590-2161	
43	15.2	76.0	663	13	US-10-027-632-187538	
44	15.2	76.0	663	13	US-10-027-632-187539	
45	15.2	76.0	663	13	US-10-027-632-187539	

ALIGNMENTS

RESULT 1
US-09-834-291-5
Sequence 5, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
CURRENT FILING DATE: 2001-08-21
PRIOR FILING DATE: 1999-10-18
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 199 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-5
Query Match 100.0% Score 20; DB 9; Length 20;
Best Local Similarity 100.0% Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACAGCCTGACACGCA 20
DB 1 GGCACAGCCTGACACGCA 20
RESULT 2
US-09-834-291-10
Sequence 10, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-10

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20
DB 1 GGACAAGCCCTGACAAGCCA 20

RESULT 3
US-09-834-291-24
Sequence 24, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-24

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACAAGCCCTGACAAGCCA 20
DB 1 GGACAAGCCCTGACAAGCCA 20

RESULT 4
US-09-834-291-32
Sequence 32, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 32
LENGTH: 266
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-32

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 266;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20
DB 160 GGACAAGCCCTGACAAGCCA 179

RESULT 5
US-09-834-291-2
Sequence 2, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 720
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-2

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 720;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20
DB 160 GGACAAGCCCTGACAAGCCA 179

RESULT 6
US-09-834-291-3
Sequence 3, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 2380

Sun Aug 3 09:03:43 2003

us-09-834-291-10.rmpb

Page 3

TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-3

Query Match 100.0%; Score 20; DB 9; Length 2380;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAAAGCCA 20
DB 1820 GGACAAAGCCTGACAAAGCCA 1839

RESULT 7
US-09-834-291-4
Sequence 4, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2827
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-4

Query Match 100.0%; Score 20; DB 9; Length 2827;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAAAGCCA 20
DB 2267 GGACAAAGCCTGACAAAGCCA 2286

RESULT 8
US-09-834-291-1
Sequence 1, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3212
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-1

Query Match 100.0%; Score 20; DB 9; Length 3212;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAAAGCCA 20
DB 2659 GGACAAAGCCTGACAAAGCCA 2678

RESULT 9
US-09-834-291-6
Sequence 6, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-6

Query Match 92.0%; Score 18.4; DB 9; Length 20;
Best Local Similarity 95.0%; Pred. No. 7.4;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAAAGCCA 20
DB 1 GGAAAGCCTGACAAAGCCA 20

RESULT 10
US-09-834-291-7
Sequence 7, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-7

Query Match 84.0%; Score 16.8; DB 9; Length 20;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAAAGCCA 20
DB 1 GGAAAGCCTGACAAAGCCA 20

RESULT 11

US-09-834-291-9
Sequence 9, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-9

Query Match 84.0%; Score 16.8; DB 9; Length 20;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACAGCCCTGACAGCCA 20
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 12
US-10-027-632-174763/C
Sequence 174763, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 174763
LENGTH: 2940917
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1).....(2940917)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174763

Query Match 84.0%; Score 16.8; DB 13; Length 2940917;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACAGCCCTGACAGCCA 20

DB 2746843 GGACAGCCCTGAATGCCA 2746824

RESULT 13
US-09-728-444-1101/C
Sequence 1101, Application US/09728444
Patent No. US20020161207A1
GENERAL INFORMATION:
APPLICANT: Friedlich, Glenn
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20020161207A1 Marine Polynucleotide Sequences
FILE REFERENCE: LEX-0100-USA
CURRENT APPLICATION NUMBER: US/09/728,444
CURRENT FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/168,360
PRIOR FILING DATE: 1999-12-01
NUMBER OF SEQ ID NOS: 1206
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1101
LENGTH: 516
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)....(516)
OTHER INFORMATION: n = A,T,C or G
US-09-728-444-1101

Query Match 82.0%; Score 16.4; DB 10; Length 516;
Best Local Similarity 94.4%; Pred. No. 74;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GACAGCCCTGACAGCC 19
DB 100 GACAGCCCTGACAGCC 83

RESULT 14
US-09-983-965-2743/C
Sequence 2743, Application US/09983965
Patent No. US20020137160A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Ningling
APPLICANT: Byatt, John C.
APPLICANT: Mathalagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 2743
LENGTH: 433
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
OTHER INFORMATION: Clone ID: 13-LIB3058-025-01-K1-D1
US-09-983-965-2743

Query Match 80.0%; Score 16; DB 10; Length 433;
Best Local Similarity 100.0%; Pred. No. 12e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CAAGCCCTGACAGCC 19

Sun Aug 3 09:03:43 2003

US-09-834-291-10.rnpb

Page 5

Db 102 CAAGCCCTGACAAGCC 87

RESULT 15

US-09-878-574-9518

Sequence 9518, Application US/09878574

Patent No. US20020110548A1

GENERAL INFORMATION:

APPLICANT: Byrum, Joseph R.

APPLICANT: La Rosa, Thomas J.

APPLICANT: Thompson, Michael D.

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(15401)B

CURRENT APPLICATION NUMBER: US/09/878,574

PRIOR FILING DATE: 2001-12-21

PRIOR FILING DATE: 1999-06-14

NUMBER OF SEQ ID NOS: 15775

SEQ ID NO 9518

LENGTH: 165

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: 701102580H1

US-09-878-574-9518

Query Match 79.0%; Score 15.8; DB 10;

Best Local Similarity 89.5%; Pred. No. 1.5e+02;

Matches 17; Conservative 0; Mismatches 2;

QY 1 GGACAAGCCCTGACAAGCC 19

Db 131 GGACAAGCTCGACAAGCC 149

Length 165;

Indels 0; Gaps 0;

Search completed: August 1, 2003, 13:36:56
Job time: 17.1861 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds
(without alignments)
10468.541 Million cell updates/sec

Title: US-09-834-291-12

Perfect score: 20
Sequence: 1 agagatgcccaactgttct 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_str:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_mus:*
33: em_hcg_mus:*
34: em_hcg_mus:*
35: em_hcg_mus:*
36: em_hcg_mus:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_hum:*
41: em_hcg_hum:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	20	100.0	20	6	AX026100	AX026100 Sequence
2	20	100.0	40	6	AX026106	AX026106 Sequence
3	20	100.0	40	6	AX026114	AX026114 Sequence
4	20	100.0	2344	6	HSCD955FR	X87625 H. sapiens C
5	20	100.0	2827	6	AX026092	AX026092 Sequence
6	20	100.0	3212	6	AX026089	AX026089 Sequence
7	20	100.0	45121	6	AX695635	AX695635 Sequence
8	20	100.0	187313	9	AL157394	AL157394 Human DNA
9	19	95.0	193566	2	AC142173	AC142173 Human DNA
10	18.4	92.0	74951	2	AL513364	AL513364 Human DNA
11	18.4	92.0	163453	9	AC099561	AC099561 Homo sapi
12	18.4	92.0	165071	2	EX005378	EX005378 Homo sapi
13	18.4	92.0	168522	2	AC024302	AC024302 Homo sapi
14	18.4	92.0	182535	2	EX323040	EX323040 Homo sapi
15	18.4	92.0	236685	2	AC084744	AC084744 Mus muscu
16	18	90.0	110000	2	AC099203	AC099203 Rattus no
17	18	90.0	215492	2	AC128141	AC128141 Rattus no
18	17.4	87.0	645	8	AY290110	AY290110 Zea mays
19	17.4	87.0	782	8	ATRS31858	ATRS31858 Arabidops
20	17.4	87.0	16157	1	AE001670	AE001670 Chlamydia
21	17.4	87.0	17395	1	AE002255	AE002255 Chlamydia
22	17.4	87.0	20563	2	AC015202	AC015202 Drosophila
23	17.4	87.0	110000	6	AR310754	AR310754 Drosophila
24	17.4	87.0	121251	5	AL591593	AL591593 Zebrafish
25	17.4	87.0	129606	2	AC018486	AC018486 Drosophila
26	17.4	87.0	147727	9	AC079160	AC079160 Homo sapi
27	17.4	87.0	160121	2	EX088586	EX088586 Homo sapi
28	17.4	87.0	163448	10	AC116579	AC116579 Mus muscu
29	17.4	87.0	170729	2	AC118688	AC118688 Mus muscu
30	17.4	87.0	175326	2	AL954772	AL954772 Dario rer
31	17.4	87.0	181053	3	AC022346	AC022346 Drosophila
32	17.4	87.0	200098	5	AL935317	AL935317 Zebrafish
33	17.4	87.0	212154	2	EX323453	EX323453 Dario rer
34	17.4	87.0	222313	2	AC094871	AC094871 Rattus no
35	17.4	87.0	235483	2	AC134161	AC134161 Rattus no
36	17.4	87.0	254946	2	AL844587	AL844587 Dario rer
37	17.4	87.0	258988	2	AC123558	AC123558 Mus muscu
38	17.4	87.0	299868	3	AC103322	AC103322 Rattus no
39	17.4	87.0	325865	1	AE003496	AE003496 Drosophila
40	17.4	87.0	325865	1	AP002548	AP002548 Chlamydia
41	17	85.0	34200	9	AB022785	AB022785 Homo sapi
42	17	85.0	37068	8	AC067938	AC067938 Neurospor
43	17	85.0	43556	8	AC067937	AC067937 Neurospor
44	17	85.0	74043	2	AC121551	AC121551 Mus muscu
45	17	85.0	96376	2	AC007626	AC007626 Homo sapi

ALIGNMENTS

RESULT 1
AX026100
LOCUS
DEFINITION
AX026100
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Sequence 12 from Patent DE19847779.
AX026100.1 GI:10187531

REFERENCE
AUTHORS
TITLE
1
Mueller-Schilling, M., Krammer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

JOURNAL

Patent: DE 19847779-C 12 03-FEB-2000;

FEATURES

DEUTSCHES KREBSFORSCH (DE)
Location/Qualifiers

SOURCE

1. .20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT

6 a 4 c 4 g 6 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AGAGATGCCCAACTGTTT 20
|||||
1 AGAGATGCCCAACTGTTT 20

Db

RESULT 2

LOCUS

AX026106 40 bp DNA linear PAT 16-SEP-2000

DEFINITION Sequence 18 from Patent DE19847779.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AGAGATGCCCAACTGTTT 20
|||||
11 AGAGATGCCCAACTGTTT 30

Db

RESULT 3

LOCUS

AX026114 40 bp DNA linear PAT 16-SEP-2000

DEFINITION Sequence 26 from Patent DE19847779.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

BASE COUNT

14 a 11 c 5 g 10 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AGAGATGCCCAACTGTTT 20
|||||
11 AGAGATGCCCAACTGTTT 30

Db

RESULT 4

LOCUS

HSCD955FR 2344 bp DNA linear PRI 05-FEB-1997

DEFINITION H.sapiens CD95 gene 5' flanking region.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

misc_signal

ACCESSION AX026092.1 GI:10187523
 VERSION
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 1984779-C 4 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)

FEATURES
 Source 1..2827
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2827;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20
 |||||
 44 AGAGATGCCCAACTGTTT 63

RESULT 6
 AX026089 3212 bp DNA linear PAT 16-SEP-2000
 LOCUS
 DEFINITION Sequence 1 from Patent DE1984779.
 AX026089
 VERSION AX026089.1 GI:10187520
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 1984779-C 1 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)

FEATURES
 Source 1..3212
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 778 a 784 c 809 g 841 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3212;
 Best Local Similarity 100.0%; Pred. No. 2.7;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20
 |||||
 44 AGAGATGCCCAACTGTTT 63

RESULT 7
 AX695635 45121 bp DNA linear PAT 31-MAR-2003
 LOCUS
 DEFINITION Sequence 1262 from Patent WO03008583.
 AX695635
 VERSION AX695635.1 GI:29418787
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Morris, D.W. and Engelhard, E.K.
 TITLE Novel compositions and methods for cancer
 JOURNAL Patent: WO 03008583-A 1262 30-JUN-2003;
 Sagres Discovery (US)

FEATURES
 Source 1..45121
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 45121;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20
 |||||
 8037 AGAGATGCCCAACTGTTT 8056

RESULT 8
 AL157394 187313 bp DNA linear PRI 22-AUG-2001
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-399019 on chromosome 10,
 complete sequence.
 AL157394
 VERSION AL157394.15 GI:15384622
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 187313)
 AUTHORS Blakey, S.
 TITLE Direct Submission
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Aug 31, 2001 this sequence version replaced gi:14161146.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: EMBL, SWISSPROT, TrEMBL, WPI, WORMPEP, information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10
 RP11-399019 is from the library RP11-11.2 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6
 This sequence is the entire insert of clone RP11-399019. The true
 left end of clone RP11-496H23 is at 166408 in this sequence. The
 true right end of clone RP11-30415 is at 18704 in this sequence.

Location/Qualifiers

```

SOURCE
1. 187313
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-399019"
/clone_1bp="RPC1-11.2"
100119. 100156
/sequence confirmed by AC015461 sequenced by WIBR."
100157. 100198
/sequence confirmed by AC015461 sequenced by WIBR."
105808. 105972
/sequence confirmed by AC015461 sequenced by WIBR."
105973. 105989
/sequence confirmed by AC015461 sequenced by WIBR."
BASE COUNT 55669 a 36398 c 36888 g 58358 t
ORIGIN
Query Match 100.0% Score 20; DB 9; Length 187313;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
Db 142012 AGAGATGCCCAACTGTTT 142031

RESULT 9
AC142173 193566 bp DNA 1linear HTG 28-MAR-2003
LOCUS Rattus norvegicus clone CH230-135C19, WORKING DRAFT SEQUENCE, 64
DEFINITION AC142173
unordered pieces.
AC142173.2 GI:29336106
HTG HTGS_PHASE1; HTGS DRAFT.
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus;
1 (bases 1 to 193566)
Muzny,D.,Maritz, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Bacca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhey,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Creg,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kows,C., Kraft,C.L., Lebow,H., Levan,Z., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,L., Louleghed,H., Lozano,R.U., Lu,X., Ma,U.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M., McNeill,T., Meenen,B., Milosavljevic,A.,
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neill,D.,
Newton,N., Nguyen,N., Norris,S., Nwackemele,O., Okonumu,G.,
Olatunbosun,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Pu,L.-L., Puzo,M., Quiroz,J., Reichlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Rosen,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojasa,A.,
Scott,G., Shatman,S., Shen,H., Shetty,J., Shwartsbeyn,A.,
Slason,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
Stavak,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseña,D.,
Waldron,L., Walker,B., Wang,J., Valas,R., Verra,V., Villaseña,D.,
Warren,R., Wei,X., White,F., Williams,G., Wilson,R., Wleczek,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,D., Yakub,S.,
Yen,J., Yoon,L., Yoon,Y., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstein,G., and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 193566)
Submitted (24-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 193566)
Worley,K.C.
Direct Submission
Submitted (28-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 28, 2003 this sequence version replaced gi:29165510.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KEON
Center clone name: CH230-135C19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 15969 bases at least Q40
Consensus quality: 16864 bases at least Q30
Consensus quality: 172461 bases at least Q20
Estimated insert size: 166404; sum-of-coverage estimation
Quality coverage: 3x in Q20 bases; sum-of-coverage estimation

```

```

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 64 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
1025: contig of 1025 bp in length
1026 1125: gap of unknown length
1126 2257: contig of 1132 bp in length
2258 2357: gap of unknown length
2358 3598: contig of 1241 bp in length
3599 3699: gap of unknown length
3699 4877: contig of 1178 bp in length
4877 4977: gap of unknown length
4977 6088: contig of 1112 bp in length
6088 8079: contig of 1891 bp in length
8079 8180: gap of unknown length
8180 9264: contig of 1084 bp in length
9264 9364: gap of unknown length
10952: contig of 1589 bp in length

```

10953 11052: gap of unknown length
 11053 12469: contig of 1417 bp in length
 12470 12569: gap of unknown length
 12570 13999: contig of 1450 bp in length
 14000 14099: gap of unknown length
 14100 15332: contig of 1233 bp in length
 15333 15432: gap of unknown length
 15433 16637: contig of 1205 bp in length
 16638 16738: gap of unknown length
 16738 18048: contig of 1311 bp in length
 18049 19368: contig of 1220 bp in length
 19369 19468: gap of unknown length
 19469 21333: contig of 1865 bp in length
 21334 21433: gap of unknown length
 21434 22762: contig of 1329 bp in length
 22763 22862: gap of unknown length
 22863 24894: contig of 2032 bp in length
 24895 24994: gap of unknown length
 24995 26516: contig of 1522 bp in length
 26517 28758: contig of 2142 bp in length
 28759 28858: gap of unknown length
 28859 30914: contig of 2056 bp in length
 30915 31014: gap of unknown length
 31015 32459: contig of 1345 bp in length
 32460 34080: gap of unknown length
 34081 34180: gap of unknown length
 34181 35284: contig of 1004 bp in length
 35285 35884: gap of unknown length
 35885 36895: contig of 1611 bp in length
 36896 36995: gap of unknown length
 36996 38112: contig of 1116 bp in length
 38112 39375: contig of 1164 bp in length
 39376 39475: gap of unknown length
 39476 41610: contig of 2135 bp in length
 41611 41710: gap of unknown length
 41710 42750: contig of 1040 bp in length
 42751 42850: gap of unknown length
 42851 45186: contig of 2336 bp in length
 45187 45286: gap of unknown length
 45287 46671: contig of 1385 bp in length
 46672 46771: gap of unknown length
 46772 48826: contig of 2055 bp in length
 48827 48926: gap of unknown length
 48927 50979: contig of 2053 bp in length
 50980 51079: gap of unknown length
 51080 52740: contig of 1701 bp in length
 52741 52880: gap of unknown length
 52881 54958: contig of 2078 bp in length
 54959 55058: gap of unknown length
 55059 56859: contig of 1801 bp in length
 56860 56959: gap of unknown length
 56960 59033: contig of 2074 bp in length
 59034 59133: gap of unknown length
 59134 61078: contig of 1945 bp in length
 61079 61178: gap of unknown length
 61179 64211: contig of 3033 bp in length
 64212 64311: gap of unknown length
 64312 66389: contig of 2078 bp in length
 66390 66489: gap of unknown length
 66490 69371: contig of 2882 bp in length
 69372 69471: gap of unknown length
 69472 71543: contig of 2072 bp in length
 71544 71643: gap of unknown length
 71644 73921: contig of 2278 bp in length
 73922 74021: gap of unknown length
 74022 77070: contig of 3049 bp in length
 77071 77170: gap of unknown length
 77171 80276: contig of 3106 bp in length
 80277 80376: gap of unknown length

Query Match 83531: contig of 3155 bp in length
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 AGAGATGCCCAACGTTT 19
 DB 167412 AGAGATGCCCAACGTTT 167430

RESULT 10
 LOCUS AL513364
 DEFINITION Human DNA sequence from clone RP11-480N10 on chromosome 1, complete sequence.
 ACCESSION AL513364
 VERSION AL513364.10 GI:16973038
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 74951)
 AUTHORS Almeida, J.
 TITLE Direct Submission
 JOURNAL Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: cloneenquiries@sanger.ac.uk
 COMMENT On Nov 16, 2001 this sequence version replaced gi:16944088. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Mp, MOPRED; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/projects/c_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr1
 RP11-480N10 is from the library RPI1-11.2 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone RP11-480N10. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true right end of clone RP11-480N10 is at 74951 in this sequence. The true left end of clone RP11-77G8 is at 42605 in this sequence. The true right end of clone RP11-54132 is at 2000 in this sequence.
 Location/Qualifiers

FEATURES


```

source
1. 74951
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-480N10"
/clone_11b="RPCI-11.2"
14660..14861
/notes="Single clone region. Sequence from reads from a
short insert library derived from a single puc
clone. Restriction digest data confirm the assembly."
48431
/notes="Tandem repeat. Forced join. Gap size estimated to
be approximately 300bp by restriction digest data."
complement(48432..48507)
/notes="Single clone region. Assembly consistent with
restriction digest data."
BASE COUNT 25517 a 14104 c 13984 g 21266 t
ORIGIN

```

```

Query Match 92.0%; Score 18.4; DB 9; Length 74951;
Best Local Similarity 95.0%; Pred. No. 34;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGAGATGCCCAACTGTTT 20
DB 51780 MGAATGCCCCAAGCTGTTT 51799

```

```

RESULT 11
AC099561 163453 bp DNA linear PRI 20-FEB-2002
LOCUS Homo sapiens chromosome 1 clone RP11-193H16, complete sequence.
DEFINITION AC099561 AL356153
AC099561.2 GI:18767669
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 163453)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saepthimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.,
Direct Submission
Unpublished
2 (bases 1 to 163453)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Submitted (16-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 163453)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saepthimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Feb 20, 2002 this sequence version replaced gi:16945998.

```

```

REFERENCE
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
JOURNAL Submitted (16-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgctg@u.washington.edu
Drafting Center: SC
Project Information
Center project name: chr-1
Center clone name: RP11-193H16 (6c0121)
Summary Statistics
Sequencing vector: plasmid, 108752, 87% of reads
Chemistry: Dye-terminator ET; 68% of reads
Chemistry: Dye-terminator Big Dye; 32% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163426 bases at least Q40

```

Consensus quality: 163453 bases at least Q30
 Consensus quality: 163453 bases at least Q20
 Insert size: 180697; sum-of-contigs
 Quality coverage: 8.2k in Q20 bases; sum-of-contigs

Overlapping Sequences:
 3' : RP11-63117 AL359971, 2000-bp overlap
 5' : RP4-722133 (UWGC:6c0558) AL365355

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8565	1949	1903	15694	15823
6	<800	6382	6533	2067	2083
10511	10321	512	<800	7215	7441
3251	3411	449	<800	1979	1986
2717	2761	2776	2753	445	<800
1023	1057	2700	2753	1292	1278
4241	4333	5132	5075	1170	1163
1831	1830	12787	12698	3837	3814
7818	7675	1103	1091	3564	3572
932	918	1317	1257	837	906
12395	12928	5920	5812	5589	5567
2809	2761	1887	1903	58	<800
183	<800	5786	5812	8040	8112
12956	12928	5002	5075	998	906
1588	1586	8584	8683	508	<800

5958	5997	9300	9335	16398	15823
1485	1456	1942	1903	4873	4875
3790	3672	529	<800	2847	2834
2032	2037	3774	3815	4216	4112
648	<800	519	<800	2429	2499
335	<800	1939	1903	2781	2834
3249	3411	471	<800	4688	4653
5628	5579	4026	3985	3544	3572
3046	3020	2288	2347	12025	11730
2433	2442	263	<800	4093	4112
2990	3020	1478	1414	2248	2274
7657	7675	7155	7408	2317	2274
1341	1321	3557	3526	5732	5567
1825	1830	2397	2347	2481	2499
143	<800	863	894	3030	3022
3099	3244	11609	11460	253	<800
3480	3411	971	894	4136	4112
9141	9053	2669	2640	1507	1495
4104	4099	1890	1903	5520	5567
976	918	8666	8683	22304	22665
115	<800	1280	1257	295	<800
3601	3672	4288	4259	1752	1743
218	<800	4227	4259	175	<800
32	<800	325	<800	3783	3814
10290	10321	2186	2178	3843	3814
5337	5272	6371	6282	1589	1562
8075	7675	859	894		
2662	2761	2630	2640		
1733	1710	2009	2031		
3006	3020	845	894		
2726	2761	7421	7408		
		377	<800		
		1595	1590		
		1467	1414		
		851	894		
		1214	1257		

FEATURES

source

Location/Qualifiers
1. 163453
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"

Query Match 92.0%; Score 18.4; DB 9; Length 163453;
Best Local Similarity 95.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 AGAATGCCCAACTGTTT 20
DB 150211 AGAATGCCCAACTGTTT 150192

RESULT 12

BX005378

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

McLay, K.

Submitted (02-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,

Cambridge, CB10 1SA, UK. E-mail enquiries: zfac@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

On Jan 4, 2003 this sequence version replaced gi:27475575.

Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: zfac@sanger.ac.uk

Project Information

Center project name: zK40H20

Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: dye-terminator; 100% of reads

Consensus quality: 161469 bases at least Q40

Consensus quality: 162178 bases at least Q30

Consensus quality: 162708 bases at least Q20

Insert size: 164371; sum-of-contigs

Insert size: 163675; 5.4% error; agarose-fp

Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality

Coverage: 6.54x in Q20 bases; agarose-fp

be preserved

1

36150: contig of 36150 bp in length

36151

36250: gap of 100 bp

36251

42442: contig of 6192 bp in length

42443

42542: gap of 100 bp

42543

75854: contig of 33312 bp in length

as soon as it is available and the accession number will

This record will be updated with the finished sequence

runs of N, but the exact sizes of the gaps are unknown.

arbitrary. Gaps between the contigs are represented as

is not known and their order in this sequence record is

* consists of 8 contigs. The true order of the pieces

* NOTE: This is a 'working draft' sequence. It currently

```

* 75855 75954: gap of 100 bp
* 75955 89626: contig of 13672 bp in length
* 89727 89726: gap of 100 bp
* 131894 131894: contig of 42168 bp in length
* 131895 131894: gap of 100 bp
* 134989 134989: contig of 2995 bp in length
* 135089 135089: gap of 100 bp
* 147854 147854: contig of 12765 bp in length
* 147855 147954: gap of 100 bp
* 147955 165071: contig of 17117 bp in length.

```

FEATURES

```

source
1. 165071
Location/Qualifiers

```

```

misc_feature
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_id="DKEY-40H20"
/clone_lib="DanioKey"
1. 36150
/note="assembly fragment:00533
fragment chain:1"
36251..42442
/note="assembly fragment:00652
fragment chain:1"
42543..75854
/note="assembly fragment:01251
fragment chain:1"
75955..89626
/note="assembly fragment:01739
fragment chain:1"
89727..131894
/note="assembly fragment:01951
fragment chain:2"
131955..134989
/note="assembly fragment:00031
fragment chain:2"
135090..147854
/note="assembly fragment:01256
fragment chain:2"
147955..165071
/note="assembly fragment:01836
fragment chain:2"
BASE COUNT 54498 a 30059 c 29325 g 50484 t 705 others
ORIGIN
Query Match 92.0%; Score 18.4; DB 2; Length 165071;
Best Local Similarity 95.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AGAGATGCCCAACTGTTT 20
Db 107313 AGAGATGCCCAACTGTTT 107332

```

```

RESULT 13
AC024302/c 168522 bp DNA linear HTG 14-MAR-2000
LOCUS AC024302
DEFINITION Homo sapiens clone RP11-23123, WORKING DRAFT SEQUENCE, 18 unordered
ACCESSION AC024302
VERSION AC024302.2 GI:7239599
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 168522)
Britten, B., Linton, L., Nussbaum, C., and Lander, E.
REFERENCE
JOURNAL
Unpublished
2 (bases 1 to 168522)
Britten, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barina, N., Beda, F., Boguslavsky, L.,
Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,

```

TITLE
JOURNAL
COMMENT

Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Donato, M., Doyle, M.,
Fenster, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goette, M., Graham, L.,
Grand, P., Grant, G., Hago, B., Heaford, A., Horton, L.,
Howland, J., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Landers, T., Laroque, K., Lehoucq, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., MacDonald, P., Margu, N., McCarthy, M.,
McEwan, P., McGurk, A., McKernan, K., McPherson, R., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., O'Leary, J.,
Peterson, K., Pierre, N., Pisan, C., Polata, V., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A.,
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and
Zody, M.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: L4376
Center clone name: 23_1_23

Sequencing Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 155724 bases at least Q40
Consensus quality: 162423 bases at least Q30
Insert size: 168000; agarose-gel
Insert size: 166822; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-gel
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1 1007: contig of 1007 bp in length
* 1008 1107: gap of 100 bp
* 1108 2229: contig of 1122 bp in length
* 2230 2339: gap of 100 bp
* 2339 4120: contig of 1791 bp in length
* 4121 4220: gap of 100 bp
* 4221 6093: contig of 1873 bp in length
* 6094 6193: gap of 100 bp
* 6194 12656: contig of 6463 bp in length
* 12657 12756: gap of 100 bp
* 12757 19312: contig of 6555 bp in length
* 19312 19411: gap of 100 bp
* 19411 24024: contig of 4613 bp in length
* 24025 24124: gap of 100 bp
* 24125 30616: contig of 6492 bp in length
* 30617 30716: gap of 100 bp
* 30717 35830: contig of 5114 bp in length
* 35831 35930: gap of 100 bp
* 35931 45340: contig of 9410 bp in length
* 45341 45440: gap of 100 bp

```

```

* 45441 55960: contig of 10520 bp in length
* 55961 gap of 100 bp
* 56061 67711: contig of 11651 bp in length
* 56061 67711: gap of 100 bp
* 67712 78627: contig of 10816 bp in length
* 67712 78627: gap of 100 bp
* 78628 91285: contig of 12558 bp in length
* 78628 91285: gap of 100 bp
* 91286 106039: contig of 14654 bp in length
* 91286 106039: gap of 100 bp
* 106040 123134: contig of 16395 bp in length
* 106040 123134: gap of 100 bp
* 123135 140131: contig of 16897 bp in length
* 123135 140131: gap of 100 bp
* 140132 168522: contig of 28291 bp in length.
* 140232 168522: Location/Qualifiers
  1..168522
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone_id="RP11-23123"
    /clone_1b="RP11-11 Human Male BAC"
  
```

```

misc_feature 1..1007
              /note="assembly_fragment"
misc_feature 1108..2229
              /note="assembly_fragment"
misc_feature 2330..4120
              /note="assembly_fragment"
misc_feature 4221..6093
              /note="assembly_fragment"
misc_feature 6194..12656
              /note="assembly_fragment"
misc_feature 12757..19311
              /note="assembly_fragment"
misc_feature 19412..24024
              /note="assembly_fragment"
misc_feature 24125..30616
              /note="assembly_fragment"
misc_feature 30717..35830
              /note="assembly_fragment"
misc_feature 35931..45340
              /note="assembly_fragment"
misc_feature 45441..55960
              /note="assembly_fragment"
              vector_side:right"
              clone_end:SP6
misc_feature 56061..67711
              /note="assembly_fragment"
misc_feature 67812..78627
              /note="assembly_fragment"
              clone_end:17
              vector_side:left"
misc_feature 78728..91285
              /note="assembly_fragment"
misc_feature 91386..106039
              /note="assembly_fragment"
misc_feature 106140..123134
              /note="assembly_fragment"
misc_feature 123235..140131
              /note="assembly_fragment"
misc_feature 140232..168522
              /note="assembly_fragment"
BASE COUNT 54289 a 31221 c 31549 g 49758 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 2; Length 168522;
Best Local Similarity 95.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 AGAGATGCCCAACTGTTT 20
|||||
DB 145086 AGAATGCCCAACTGTTT 145067

```

```

RESULT 14
BX323040 182635 bp DNA linear HTG 24-APR-2003
LOCUS Homo sapiens chromosome 1 clone RP11-455622, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
ACCESSION BX323040 4 GI:30140477
VERSION HTG; HTGS_PHASE1; HTGS_CANCELLED.
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens [Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          Diastelle H.
          1 (bases 1 to 182635)]
REFERENCE Direct Submission
          Submitted (23-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
          On Apr 25, 2003 this sequence version replaced gi:30026997.
JOURNAL ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: humquerry@sanger.ac.uk
          Project Information
          ----- Summary Statistics
          Assembly program: XGAP4; version 4.5
          Sequencing vector: plasmid; L08752; 100% of reads
          Chemistry: Dye-terminator; 75% of reads
          Chemistry: Dye-terminator; ET-amersham; 4% of reads
          Chemistry: Dye-terminator; Big Dye; 19% of reads
          Consensus quality: 181801 bases at least Q40
          Consensus quality: 182063 bases at least Q30
          Consensus quality: 182266 bases at least Q20
          Insert size: 182435; sum-of-contigs
          Insert size: 177453; 5.1% error; agarose-fp
          Quality coverage: 24.85x in Q20 bases; sum-of-contigs Quality
          coverage: 29.90x in Q20 bases; agarose-fp
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 3 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          1 178116: contig of 178116 bp in length
          * 178117 178216: gap of 100 bp
          * 178217 180249: contig of 2033 bp in length
          * 180250 180349: gap of 100 bp
          * 180350 182635: contig of 2286 bp in length.
          Location/Qualifiers
          1..182635
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="RP11-455622"
            /clone_1b="RP11-11.2"
            1..178116
              /note="assembly_fragment:00629"
              178217..180249
                /note="assembly_fragment:01241"
                180350..182635
                  /note="assembly_fragment:03792"
                  /note="assembly_fragment:03792"
BASE COUNT 59564 a 36041 c 34489 g 52341 t 200 others
ORIGIN
Query Match 92.0%; Score 18.4; DB 2; Length 182635;
Best Local Similarity 95.0%; Pred. No. 37;

```

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATCCCCAACTGTTT 20
Db 121795 AGAATGCCCAACTGTTT 121814

RESULT 15
AC084744/c
LOCUS
DEFINITION
AC084744 236685 bp DNA linear HTG 03-FEB-2001
Mus musculus clone RP23-15A13, WORKING DRAFT SEQUENCE, 35 unordered
pieces
AC084744
AC084744 GI:12658010
HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
McCombie, W.R., Baker, J.P., Bahret, A., Bal, H., Balija, V.,
Dedhia, N.N., de la Bastide, M., Huang, E.N., King, L., Kirchhoff, K.A.,
Rodriguez, S., Santos, L., Shah, R.S., Spiegel, L.A., Toch, K., Vill, M.D.,
and Zlatavern, T.
Mouse Genomic Sequence
Unpublished
2 (bases 1 to 236685)
McCombie, W.R.
Direct Submission
Submitted (11-NOV-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
On Feb 3, 2001 this sequence version replaced gi:11276152.
----- Genome Center
Laboratory: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Center code: CSHL
Web site: http://www.cshl.org/geneseq
Contact: mcombie@cshl.org
----- Project Information
Center project name: RP23-15A13
Center clone name: RP23-15A13

* NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
22323: contig of 22323 bp in length
22324
22432: gap of unknown length
22433
42140: contig of 19708 bp in length
42141
42249: gap of unknown length
42250
57893: contig of 15644 bp in length
57894
58002: gap of unknown length
72097
72098
72206: contig of 14095 bp in length
72207
85432: gap of unknown length
85433
85541: gap of unknown length
85542
98091: contig of 13226 bp in length
98092
98200: gap of unknown length
98201
110627
110628: gap of unknown length
110629
110735: gap of unknown length
110736
119883: contig of 9148 bp in length
119884
119991: gap of unknown length
119992
128840: contig of 8849 bp in length
128841
128948: gap of unknown length
128949
137470: contig of 8522 bp in length
137471
137579
145833: contig of 8255 bp in length

145834 145941: gap of unknown length
145942 152662: contig of 6721 bp in length
152663 152770: gap of unknown length
152771 158664: contig of 5894 bp in length
158665 158772: gap of unknown length
158773 164620: contig of 5848 bp in length
164621 164728: gap of unknown length
164729 170566: contig of 5838 bp in length
170567 170674: gap of unknown length
170675 172684: contig of 5590 bp in length
172685 176373: gap of unknown length
176374 181545: contig of 5173 bp in length
181546 181653: gap of unknown length
181654 186814: contig of 5161 bp in length
186815 186922: gap of unknown length
186923 191171: contig of 4249 bp in length
191172 191279: gap of unknown length
191280 195507: contig of 4228 bp in length
195508 195508: gap of unknown length
195509 195616: gap of unknown length
195617 199208: contig of 3593 bp in length
199209 199316: gap of unknown length
199317 202691: contig of 3375 bp in length
202692 202799: gap of unknown length
202800 206145: contig of 3346 bp in length
206146 206253: gap of unknown length
206254 206254: gap of unknown length
206255 209554: contig of 3301 bp in length
209556 209662: gap of unknown length
209663 212567: contig of 2905 bp in length
212568 212675: gap of unknown length
212676 215151: contig of 2840 bp in length
215152 215623: gap of unknown length
215624 218450: contig of 2827 bp in length
218451 218558: gap of unknown length
218559 221350: contig of 2792 bp in length
221351 221458: gap of unknown length
221459 224193: contig of 2735 bp in length
224194 224301: gap of unknown length
224302 226669: contig of 2368 bp in length
226670 226777: gap of unknown length
226778 229025: contig of 2248 bp in length
229026 229133: gap of unknown length
229134 231217: contig of 2084 bp in length
231218 231325: gap of unknown length
231326 233333: contig of 2008 bp in length
233334 233441: gap of unknown length
233442 235387: contig of 1946 bp in length
235388 235495: gap of unknown length
235496 236685: contig of 1190 bp in length.

FEATURES
source
1. 236685
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-15A13"

BASE COUNT 58998 a 57976 c 57568 g 58324 t 3819 others

ORIGIN

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 2; Length 236685;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATCCCCAACTGTTT 20
Db 236285 AGAATGCCCAACTGTTT 236266

Search completed: August 1, 2003, 17:32:53
Job time : 82.1573 secs

Sun Aug 3 09:03:53 2003

us-09-834-291-12.fst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds
(without alignments)
7098.748 Million cell updates/sec

Title: US-09-834-291-12

Perfect score: 20

Sequence: 1 agagatgcccaactgttt 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbv:*
5: em_estbv:*
6: em_estbv:*
7: em_estbv:*
8: em_estbv:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estbm:*
16: em_estbm:*
17: em_estbm:*
18: em_estbm:*
19: em_estbm:*
20: em_estbm:*
21: em_estbm:*
22: em_estbm:*
23: em_estbm:*
24: em_estbm:*
25: em_estbm:*
26: em_estbm:*
27: em_estbm:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	90.0	539	9	AL922459
2	17.4	87.0	282	10	BB573416
3	17.4	87.0	475	12	BM420516
4	17.4	87.0	635	9	AL963961

C	5	17.4	87.0	677	10	BG410123	BG410123	510-8-A10
C	6	17.4	87.0	735	10	BB630820	BB630820	BB630820
C	7	17.4	87.0	796	28	BH393256	BH393256	AG-ND-165
C	8	17.2	86.0	1201	13	EX422214	EX422214	EX422214
C	9	17	85.0	1035	9	AL575246	AL575246	AL575246
C	10	17	85.0	1087	9	AL551725	AL551725	AL551725
C	11	16.8	84.0	200	28	A2766513	A2766513	1M0564H07
C	12	16.8	84.0	291	9	AA400303	AA400303	zue4a03.x
C	13	16.8	84.0	302	10	BF835617	BF835617	RC4-HT089
C	14	16.8	84.0	314	10	BF994308	BF994308	CM2-GN016
C	15	16.8	84.0	322	28	A2719811	A2719811	RPC1-24-1
C	16	16.8	84.0	326	10	BB508020	BB508020	BB508020
C	17	16.8	84.0	335	14	D60325	D60325	HUM099E06A
C	18	16.8	84.0	377	10	AM879474	AM879474	PMO-OT001
C	19	16.8	84.0	385	10	BF994425	BF994425	CM2-GN016
C	20	16.8	84.0	394	28	A2231468	A2231468	RPC1-23-6
C	21	16.8	84.0	411	28	AO173492	AO173492	HS-3193.A
C	22	16.8	84.0	466	28	BZ118730	BZ118730	CH230-460
C	23	16.8	84.0	496	10	CA332861	CA332861	bae27c11.
C	24	16.8	84.0	522	14	CA332861	CA332861	bae27c11.
C	25	16.8	84.0	522	14	CA332861	CA332861	bae27c11.
C	26	16.8	84.0	535	9	AM704890	AM704890	1E83907.Y
C	27	16.8	84.0	546	28	AO992080	AO992080	sk40h01.Y
C	28	16.8	84.0	563	10	BF191086	BF191086	nbe0080C
C	29	16.8	84.0	578	28	AZ909829	AZ909829	237803.MA
C	30	16.8	84.0	611	29	BZ294769	BZ294769	RCPI-24-2
C	31	16.8	84.0	622	28	BH762728	BH762728	CGI081.F1
C	32	16.8	84.0	634	28	CC093693	CC093693	EMBAC30F
C	33	16.8	84.0	667	28	BZ02104	BZ02104	CSU-R34.1
C	34	16.8	84.0	667	28	AG037563	AG037563	cee36e03.
C	35	16.8	84.0	673	29	BZ920990	BZ920990	pan trogl
C	36	16.8	84.0	676	14	BY727273	BY727273	CH240.118
C	37	16.8	84.0	684	13	EX101371	EX101371	BY727273
C	38	16.8	84.0	688	12	BH096294	BH096294	EX101371
C	39	16.8	84.0	693	13	BH388813	BH388813	EX101371
C	40	16.8	84.0	698	29	BX184556	BX184556	EX101371
C	41	16.8	84.0	700	28	BZ074128	BZ074128	EX101371
C	42	16.8	84.0	742	29	BX205069	BX205069	EX101371
C	43	16.8	84.0	748	28	BH988183	BH988183	EX101371
C	44	16.8	84.0	783	29	AG094259	AG094259	pan trogl
C	45	16.8	84.0	783	29	AG094259	AG094259	pan trogl

ALIGNMENTS

RESULT 1
LOCUS AL922459 539 bp mRNA linear EST 18-SRP-2002
DEFINITION AL922459 PUR-Z1+Z2 Danio rerio cDNA clone 111-E05-2, mRNA sequence.
ACCESSION AL922459
VERSION AL922459.1 GI:23189039
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
REFERENCE Lee, S., Ruan, H., Ma, W., Xu, M., Lo, J., He, Y., Liu, F., Sun, A., Wen, Z.
AUTHORS 1 (bases 1 to 539)
TITLE 15,000 unique zebrafish EST clusters from two cDNA libraries
JOURNAL Unpublished
COMMENT Contact: Peng J
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengj@imcb.a-star.edu.sg
Clone requests: pengj@imcb.a-star.edu.sg
Location/Qualifiers
1..539
/organism="Danio rerio"
/mol_type="mRNA"

RESULT 2	BB573416/c	LOCUS	DEFINITION	VERSION	ACCESSION	KEYWORDS	SOURCE
	BB573416	282 bp	mRNA	linear	EST 30-NOV-2000		
	BB573416	musculus full-length enriched, adult male testis (DH10B)					
	BB573416	musculus CDNA clone 4930501C15 5', mRNA sequence.					
	BB573416.1	GI:11469960					
	EST.						
	Mus musculus (house mouse)						

Mus musculus (house mouse)	Mus musculus	ORGANISM
----------------------------	--------------	----------

REFERENCE AUTHORS

TITLE Havesha-ni Yashiki ni Kiteba
JOURNAL *Journal of Japanese Linguistics*
COMMENT
Contact: Yoshihide Havesha-ni Yashiki ni Kiteba
Unpublished
ESTs (Aizawa, K. et al. 2000)

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
1-7-22 Sengen-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermotranscription and thermostabilization of thermophilic enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2): 520-524 (1998)
Itoh, M., Kitsunai, T., Akiyama, J., Shibata, K., Iwama, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5): 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)
Please visit our web site (<http://genome.ritc.riken.go.jp/>) for further details.

ATTURES
SOURCE

```

/organism="Mus musculus
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="493050IC15"

```

```

/sex="male"
/classic_type="testis"
/dev stage="adult"
/lab_host="DH10B"
/clone_idb="RIKEN full-length enriched, adult male testis
(DH10B)"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGAGATCCAGAGCCTCTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using triethanol thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified plasmidic KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI."

```

	Query Match	87.0%; Score 17.4; DB 10;	length 282;
	Best Local Similarity	94.7%; Pred. No. 3e+02;	
Matches	18; Conservative	0; Mismatches	1; Indels 0; Gaps 0.
QY	2 GAGATGCCCAACTGTTT	20	
Dd	149 GAGATGCCCAACTGTCTT	131	

RESULT 3	LOCUS	DEFINITION	ACCSSION
BM420516	475 bp	mRNA linear EST 28-JAN-2002	
U013C07	Oryza sativa mature leaf library induced by M.grisea	Oryza	
BM420516		cDNA clone U013C07, mRNA sequence.	

VERSION
 EM420516.1 GI:18387318
 EST.
 KEYWORDS
 ORYZA SATIVA
 SOURCE
 ORYZA SATIVA
 ORGANISM
 Oryza sativa
 Bukayocsa, Viridiplantae; Streptophyta; Embryophyta; Spermatophyta, Magnoliophyta, Liliopsida, Euphorbiales, Euphorbiaceae, Oryzoideae, Oryza.
 1 (bases 1 to 475)
 REFERENCE
 Dong, H. T., Li, D. B., Zhuang, Y. T.

TITLE
JOURNAL
COMMENT

H.F., Jiang, Y.X., Yu, F.C., Gao, Q.K. and Lou, Y.C.
A Gene Expression Screen in *Oryza sativa*
Unpublished
Contact: Dong HT

FEATURES
source
Laboratory of Functional Genetics
Bio-technology Institute of Zhejiang University
Kaixuan Road 268# Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: htdong@zjuem.zju.edu.cn
Seq primer: M13 forward primer.
Location/Qualifiers
1. 475

```
/organism="Oryza sativa"  
/mol_type="mRNA"  
/db_xref="taxon:4530"  
/clone="U013C07"
```

/tissue_type="leaf"	
/dev_stage="Mature stage"	
/clone_idb="Oryza sativa mature leaf library induced by	
M.grisea"	
/note="Vector: pSport2"	
88 c	109 g
140 a	137 t
	1.others

Sun Aug 3 09:03:53 2003

us-09-834-291-12.rst

Page 3

Query Match 87.0%; Score 17.4; DB 12; Length 475;
Best Local Similarity 94.7%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1;

QY 1 AGAGTGGCCCAACTGTTT 19
178 AGAGTGGCCCAACTGTTT 216

RESULT 4
AL963961/c 635 bp mRNA linear EST 27-NOV-2002
LOCUS AL963961 XGC-gastrula silurana tropicalis cDNA clone TGA511h03 5',
DEFINITION mRNA sequence.

ACCESSION AL963961 GI:25787556
VERSION AL963961
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Anura; Mesobatrachia; Pipidae; Pipidae;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopus laevis; Silurana.

REFERENCE 1 (bases 1 to 635) Croxson, M.D.R., Zorn, A.M. and Rogers, J.
AUTHORS Taylor, R., Ashurst, J.U., Croxson, M.D.R., Zorn, A.M. and Rogers, J.
TITLE Sanger Xenopus tropicalis EST project 2002
JOURNAL Unpublished
COMMENT Contact: Taylor R
Sanger Centre
Cambridge, CB10 1SA, UK
Email: tropesanger.ac.uk
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPOCALIS SEQUENCE ID: TGA511h03.p1ksp6
TROPOCALIS SEQUENCE ID: TGA511h03.p1ksp6
Sequencing primer: Sp6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..635
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TGA511h03"
/dev_stage="gastrula (stages 10.5-13 mixed)"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-gastrula"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end." end.. 110 c 197 g 138 t

FEATURES
source

BASE COUNT 190 a 110 c 197 g 138 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 635;
Best Local Similarity 94.7%; Pred. No. 4.6e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1;

QY 2 GAGATGCCCAACTGTTT 20
178 GAGATGCCCAACTGTTT 160

RESULT 5
BG410123/c 677 bp mRNA linear EST 01-APR-2001
LOCUS BG410123 S10-8-A10 Stage 10+ Gastrula library Xenopus laevis cDNA 5' similar
DEFINITION to Xenopus EST, mRNA sequence.

ACCESSION BG410123 GI:13506129
VERSION BG410123
KEYWORDS EST.
Xenopus laevis (African clawed frog)
Xenopus laevis
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 677) Sczyrba, A., Pun, J., Bekiranov, S.,
Altman, C.R., and Brivanlou, A.H.
Gaasterland, T. and Brivanlou, A.H.
Microarray-based analysis of early development in Xenopus laevis
Dev. Biol. 236 (1), 64-75 (2001)

JOURNAL MEDLINE 21349872
PUBMED 11456444
Contact: Brivanlou, AH
Laboratory of Molecular Vertebrate Embryology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Tel: 212 327 8684
Fax: 212 327 8685
Plate: S10-8 row: A column: 1
Seq primer: SP6-22 5' ctt gat tta ggt gac act ata g 3'.
Location/Qualifiers
1..677
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/dev_stage="10 - 10.5"
/lab_host="DH5alpha"
/clone_lib="Stage 10+ Gastrula library"
/note="Vector: pDH105/CS2+; Site 1: Sal I; Site 2: Not I;
Weinstein, D.C., Honore, E., and Hammett-Brivanlou, A. (1997)
Epidermal induction and inhibition of neural fate by
transcription factor 4A111. Development 124,
4235-4242."

FEATURES
source

BASE COUNT 197 a 141 c 190 g 149 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 677;
Best Local Similarity 94.7%; Pred. No. 4.8e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1;

QY 2 GAGATGCCCAACTGTTT 20
46 GAGATGCCCAACTGTTT 28

RESULT 6
BB630820/c 735 bp mRNA linear EST 26-OCT-2001
LOCUS BB630820 RIKEN full-length enriched, 16 days neonate thymus Mus
DEFINITION musculus cDNA clone A130029G22 5', mRNA sequence.

ACCESSION BB630820 GI:16467689
VERSION BB630820
KEYWORDS EST.
Mus musculus (house mouse)
SOURCE RIKEN Mouse ESTs (Atakawa, T., et al. 2001)
ORGANISM Mus musculus Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 735) Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Arakawa, T., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda
Hiramoto, K., Koyama, T., Miyazaki, A., Nomura, K., Ohno, M., Sasaki
, M., Koyama, S., Matsuyama, T., Miyazaki, A., Sakai, K., Sano, H., Sasaki
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki
, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Atakawa, T., et al. 2001)

JOURNAL MEDLINE 21349872
PUBMED 11456444
Contact: Yoshinaka Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saitama-cho, Tsukuba, Ibaraki, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.res.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itohi, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

Wagui, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., S., Kawai, J., Yoneeda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, Hayashizaki, Y., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamakawa, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

Location/Qualifiers

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A130029G2"
/tissue_type="thymus"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 16 days neonate
thymus"
note="Site 1: SalI, Site 2: XbaI"

```

Source

Anopheles gambiae (African malaria mosquito)
 Anopheles gambiae
 Bursaria
 Neoptera, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota,
 Anopheles, Endopterygota, Diptera, Nematocera, Culicoidae;
 1 (baes 1 to 796)
 Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J.,
 'C', Huff, E.R., Cartile, D.L., Black, K., Zhang, H.-B., Gardner, M.,
 and Collins, F.H.
 Construction of a BAC library and generation of BAC end
 sequence-tagged connectors for genome sequencing of BAC end
 malaria mosquito Anopheles gambiae
 Mol. Genet. Genomics 268 (6), 720-728 (2003)
 12542063
 12655398
 Other GSSs: AG-ND-15535, TR
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: b.loftus@tigr.org
 This clone is from an A. gambiae BAC library (ND-TM) provided by
 F.H. Collins and sequenced by The Institute for Genomic Research
 (TIGR). The BAC library was generated from A. gambiae PEST strain
 DNA. All DNA was extracted from newly hatched first instar larvae
 to minimize the inclusion of DNA from microorganisms that inhabit
 the gut. The DNA is derived from mixed sexes of larvae. The BAC
 library was constructed at Texas A&M University BAC Center
 University, College Station, Texas 77843-2123, USA using a HindIII
 partial digest.
 Seq primer: M13 For
 Class: BAC ends

Location/Qualifiers
1. 796

```

/organism="Aerobacter gambiae"
/mol_type="genomic DNA"
/strain="pSGC1m1 DNA"
/db_xref="taxon:7165"
/clone="AG-ND-16575"
/clone_1fb="ND-16575"
/notes="vector: pGEM3, Site 1: HindIII"
BASE COUNT      227 a
                   164 c
                   178 g
                   179 t

```

227 a
164 c
178 g
237 k

```

Query Match      87.0%;   Score 17.4;   DB 28;   Length 796;
Best Local Similarity 94.7%;   Pred. No. 5,2e+02;
Matches 18;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;
      1 AGGATGCCCAACTGTTT 19
      799
      DB
      Oy

```

AGATGCGGAAATTCTT

Accession	Definition	Accession	Definition
BX422214	Homo sapiens	BX422214	Homo sapiens
CSDDMO02Y119	5-PRIME, mRNA sequence.	CSDDMO02Y119	5-PRIME, mRNA sequence.
BX422214	Homo sapiens	BX422214	Homo sapiens

214

KEYWORDS
SOURCE
ORGANISM
EST.
GI:30638495
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1201)
H.W.B., Gruber, C., Teraoka, Y.

Gruber, C., Jessee, J. and Polayev, N.

variables and normalization

c: genoscope

BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5162.f For
 more information about this cluster, see:
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DM002AE10QPLcluster=5162.f. Contact:
 Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID: CS0DM002AE10QPL.

FEATURES

source
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM002Y119"
 /issue_type="FETAL LIVER"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL LIVER"
 /note="Organ: liver; Vector: PCWVSPT 6; 1st strand cDNA
 was primed with a NotI-oligo (dt) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 EcoRV sites of the PCWVSPT 6
 vector. Library was not normalized."
 Location/Qualifiers

BASE COUNT 306 a 331 c 289 g 198 t 77 others

ORIGIN

Query Match 86.0%; Score 17.2; DB 13; Length 1201;
 Best Local Similarity 85.0%; Pred. No. 8.2e+02;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 AGAGATGCCCAACTGTTT 20
 |||||
 1128 AGGATBCCCAAYGTGTTT 1109

Db 1128 AGGATBCCCAAYGTGTTT 1109

RESULT 9
 AL575246 1035 bp mRNA linear EST 01-JUN-2003
 LOCUS AL575246 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION
 AL575246 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

ACCESSION
 AL575246 GI:31313554

VERSION
 AL575246.2 GI:31313554

KEYWORDS
 EST.

SOURCE
 Homo sapiens (human)

ORGANISM
 Homo sapiens

REFERENCE
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

AUTHORS
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE
 Full-length cDNA libraries and normalization.

JOURNAL
 Unpublished

COMMENT
 On Feb 16, 2001 this sequence version replaced gi:12936226.

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID: CS0DM002AE10QPL.

Location/Qualifiers

1. 1035

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DM002Y119"

/issue_type="FETAL LIVER"

/dev_stage="fetal"

BASE COUNT 224 a 261 c 166 g 248 t 136 others

ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 1035;
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 GATGCCCAACTGTTT 20
 |||||
 201 GATGCCCAACTGTTT 217

Db 201 GATGCCCAACTGTTT 217

RESULT 10
 AL551725 1087 bp mRNA linear EST 31-MAY-2003
 LOCUS AL551725 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION
 AL551725 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

ACCESSION
 AL551725 GI:31273541

VERSION
 AL551725.2 GI:31273541

KEYWORDS
 EST.

SOURCE
 Homo sapiens (human)

ORGANISM
 Homo sapiens

REFERENCE
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

AUTHORS
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE
 Full-length cDNA libraries and normalization.

JOURNAL
 Unpublished

COMMENT
 On Feb 15, 2001 this sequence version replaced gi:12889952.

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID: CS0DM002AE10QPL.

Location/Qualifiers

1. 1087

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DM002Y118"

/issue_type="FETAL LIVER"

/dev_stage="fetal"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoRV sites of the PCWVSPT 6
 vector. Library was normalized."

Location/Qualifiers

1. 1087

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DM002Y118"

/issue_type="FETAL LIVER"

/dev_stage="fetal"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoRV sites of the PCWVSPT 6
 vector. Library was normalized."

ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 1035;
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 GATGCCCAACTGTTT 20
 |||||
 201 GATGCCCAACTGTTT 217

Db 201 GATGCCCAACTGTTT 217

RESULT 11
 A2766513 200 bp DNA linear GSS 16-FEB-2001
 LOCUS A2766513 Mouse 10kb plasmid UNGC1M library Mus musculus genomic

DEFINITION
 A2766513 Mouse 10kb plasmid UNGC1M library Mus musculus genomic

ACCESSION
 A2766513 GI:12883665

VERSION
 A2766513.1 GI:12883665

KEYWORDS
 Mus musculus (house mouse)

SOURCE
 Mus musculus (house mouse)

ORGANISM
 Mus musculus (house mouse)

REFERENCE
 Mue musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Primates; Catarrhini; Homiidae; Mus.

AUTHORS
 Mue musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Primates; Catarrhini; Homiidae; Mus.

TITLE
 Full-length cDNA libraries and normalization.

JOURNAL
 Unpublished

COMMENT
 On Feb 15, 2001 this sequence version replaced gi:12889952.

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID: CS0DM002AE10QPL.

Location/Qualifiers

1. 1087

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DM002Y118"

/issue_type="FETAL LIVER"

/dev_stage="fetal"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoRV sites of the PCWVSPT 6
 vector. Library was normalized."

Location/Qualifiers

1. 1087

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DM002Y118"

/issue_type="FETAL LIVER"

/dev_stage="fetal"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoRV sites of the PCWVSPT 6
 vector. Library was normalized."

REFERENCE
AUTHORS

1 (bases 1 to 200)

TITLE

JOURNAL

COMMENT

Dum, D., Aoyagi, A., Barber, M., Beacoin, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, F., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingy, A., von Niederhausen, A. and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished
Contact: Robert B. Weis
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddum@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Plate: 0564 row: H column: 07
Seq primer: CCGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 200.
Location/Qualifiers
1..200
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UTGCM0564H07"
/sex="male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UGCM library"
/note="Vector: PWD42nv, Purified genomic DNA from M. musculus C57BL/6J (male). The vector was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g14732114|g14732114|P12972.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

63 a 32 c 35 g 70 t

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 28; Length 200;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY

1 AGAGATGCCCAACTGTTT 20

DB

137 AGAATGCCCAACTGTTT 118

RESULT 12
LOCUS AA400303 291 bp mRNA linear EST 16-MAY-1997
DEFINITION z64a03.r1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:742732
ACCESSION AA400303
VERSION AA400303.1 GI:2054192
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

1 (bases 1 to 291)

TITLE

JOURNAL

Unpublished

COMMENT

Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through INM; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m3 rev2 87 from Amersham
High quality sequence stop: 205.
Location/Qualifiers
1..291
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5929684"
/db_xref="taxon:9606"
/clone="IMAGE:742732"
/sex="male"
/lab_host="DH10B"
/clone_1lb="Soares testis NHT"
/note="Vector: pRTT3-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATGTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

80 a 62 c 46 g 103 t

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 9; Length 291;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY

1 AGAGATGCCCAACTGTTT 20

DB

25 AGAGTGCCCAACTGCTT 6

RESULT 13
LOCUS BF835617 302 bp mRNA linear EST 13-JAN-2001
DEFINITION RC4-HT0890-151100-017-e11 HT0890 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF835617
VERSION BF835617.1 GI:12186834
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 302)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Naga, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. U., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE 20202663
PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

FEATURES

source

Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC4&l2=RC4-NT0890-151100-017-e11&l3=2000-11-15&l4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 302.
Location/Qualifiers

BASE COUNT

76 a 54 c 65 g 107 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 302;
Best Local Similarity 90.0%; Pred. No. 6.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20
Db 105 AGAATGCCCAACTGTTT 124

RESULT 14
BF994308 314 bp mRNA linear EST 23-JAN-2001

LOCUS CM2-GN0165-021100-487-803 GN0165 Homo sapiens cDNA, mRNA sequence.

DEFINITION BF994308
VERSION BF994308.1 GI:12400631

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 314)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bate, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.O.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

TITLE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE 20202663

COMMENT 10737800

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome

FEATURES

source

Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=CM2&l2=CM2-GN0165-021100-487-803&l3=2000-11-02&l4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 143.
Location/Qualifiers

BASE COUNT

71 a 75 c 74 g 94 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 314;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20
Db 150 AGACATGCCCAACTGTTT 169

RESULT 15

AZ719811 322 bp DNA linear GSS 24-JAN-2001

LOCUS RPCT-24-161H14.TV RPCT-24 Mus musculus genomic clone RPCT-24-161H14
DEFINITION , genomic survey sequence.

ACCESSION AZ719811
VERSION AZ719811.1 GI:12460860

KEYWORDS GSS.
Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 322)
Zhao, S., Nieman, W., Malek, J., Shatsman, S., Aknret, B., Levins, M.,
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,
Russell, D., de Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from library RPCT-24
Unpublished

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: shazho@tigr.org

Clones are derived from the mouse BAC library RPCT-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPRC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
pages: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 161 row: H column: 14
Seq primer: TV

Class: BAC ends.
Location/Qualifiers

1..322
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCT-24-161H14"
/sex="Male"

FEATURES

source

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

```
/cell type="Spleen/Brain"
/clone lib="Rpci-24"
/notes="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
Rpci-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      69 a      81 c      57 g      115 t
ORIGIN

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 28; Length 322;
Matches 18; Conservative 0; Pred. No. 6.6e+02;
Mismatch 2; Indels 0; Gaps 0;

QY      1 AGAGATGCCCAACTGTTT 20
        |||||
        91 AGAATGCCCAAGTGT 110

Db
```

Search completed: August 1, 2003, 20:43:03
Job time : 73.4754 secs

Sun Aug 3 09:03:48 2003

us-09-834-291-12.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds
(without alignments)
8059.612 Million cell updates/sec

Title: US-09-834-291-12

Perfect score: 20

Sequence: 1 agagatgcccaactgttt 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	20	100.0	20	21	AAZ88701	Human CD95 receptor
2	17.4	87.0	732	24	AAZ877493	Frog embryonic gen
3	17.4	87.0	1230025	20	AAZ91990	Nucleotide sequenc
4	16.8	84.0	234	23	AAZ51323	Enterococcus faeca
5	16.8	84.0	306	23	AAZ53070	Human secreted pro
6	16.8	84.0	338	21	AAZ18125	Bacillus lichenifo
7	16.8	84.0	368	24	ABK74161	Human secreted pro
8	16.8	84.0	409	24	AAZ45854	

C	9	16.8	84.0	661	16	AAZ03501	Papilloma virus ma
C	10	16.8	84.0	1569	24	AAZ45853	Human secreted pro
C	11	16.8	84.0	2008	19	AAZ92927	Human RAD54 nucle
C	12	16.8	84.0	2185	24	AAZ45822	Human secreted pro
C	13	16.8	84.0	3001	21	AAZ51750	Chromosome 13q31-q
C	14	16.8	84.0	9821	22	AAZ03245	Human reproductive
C	15	16.8	84.0	9824	22	AAZ03246	Human reproductive
C	16	16.8	84.0	20072	20	AAZ13026	Enterococcus faeca
C	17	16.8	84.0	20072	20	AAZ13026	Enterococcus faeca
C	18	16.8	84.0	43795	21	AAZ92583	Human DNZ genomic
C	19	16.8	84.0	349980	22	AAZ91123	Pyrococcus abyssi
C	20	16.4	82.0	396	25	AAZ99123	Human EST-derived
C	21	16.4	82.0	2656	25	AAZ24805	Brassica napus relat
C	22	15.8	79.0	65	24	AAZ29100	Human prostate exp
C	23	15.8	79.0	364	23	AAZ02725	Human prostate exp
C	24	15.8	79.0	416	23	AAZ11894	Human prostate exp
C	25	15.8	79.0	440	23	AAZ21752	Human prostate exp
C	26	15.8	79.0	440	23	AAZ27958	Human prostate exp
C	27	15.8	79.0	440	23	AAZ27958	Human prostate exp
C	28	15.8	79.0	440	23	AAZ28790	Human prostate exp
C	29	15.8	79.0	451	23	AAZ33039	Human prostate exp
C	30	15.8	79.0	451	23	AAZ41964	Human prostate exp
C	31	15.8	79.0	451	23	AAZ43761	Human prostate exp
C	32	15.8	79.0	454	23	AAZ13799	Human prostate exp
C	33	15.8	79.0	458	23	AAZ04630	Human prostate exp
C	34	15.8	79.0	482	23	AAZ44912	Human secreted pro
C	35	15.8	79.0	714	24	AAZ08989	Human secreted pro
C	36	15.8	79.0	1157	22	AAZ63808	Secreted protein-e
C	37	15.8	79.0	1157	25	AAZ73410	Human secreted pro
C	38	15.8	79.0	1157	25	AAZ67016	Human secreted pro
C	39	15.8	79.0	1193	21	AAZ49395	Human six transmem
C	40	15.8	79.0	1193	22	AAZ07067	Human prostate exp
C	41	15.8	79.0	1195	22	AAZ64160	Human prostate exp
C	42	15.8	79.0	1195	22	AAZ93924	P789P full length
C	43	15.8	79.0	1195	24	AAZ92128	Prostate cancer-as
C	44	15.8	79.0	1195	24	AAZ95531	Human P789P cDNA s
C	45	15.8	79.0	1195	25	AAZ59968	Prostate cancer th

ALIGNMENTS

RESULT 1	AAZ88701	standard; DNA; 20 BP.
ID	AAZ88701	
AC	AAZ88701	
XX	11-MAY-2000	(first entry)
DT	XX	Human CD95 receptor promoter DNA p53 binding fragment #1.
XX	XX	Human CD95 receptor; human; screening; apoptosis-modulation;
DE	XX	p53; CD95 receptor; human; screening; apoptosis-modulation;
XX	XX	Cancer chemotherapy; ss.
KM	XX	Homo sapiens.
XX	XX	
OS	XX	
XX	XX	DE19847779-Cl.
PN	XX	
XX	XX	03-FEB-2000.
PD	XX	
XX	XX	16-OCT-1998; 98DE-1047779.
PF	XX	
XX	XX	16-OCT-1998; 98DE-1047779.
PR	XX	
XX	XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA	XX	
XX	XX	Kramer P, Mueller-Schilling M, Oren M;
PI	XX	WPI; 2000-162245/15.
XX	XX	Novel receptor DNA useful for identifying apoptosis-modulating
DR	XX	substances potentially useful for cancer chemotherapy
XX	XX	
PT	XX	

XX Claim 2; Fig 5; 12pp; German.
 XX
 CC This invention describes a novel p53-binding region of a human CD95
 CC receptor DNA molecule. The p53-binding region, or a vector containing
 CC it, can be used to screen for apoptosis-modulating substances
 CC potentially useful for cancer chemotherapy. This sequence represents a
 CC fragment of the human CD95 receptor promoter which is capable of
 CC binding p53.
 CC

SO Sequence 20 BP; 6 A; 4 C; 4 G; 6 T; 0 other;
 Query Match
 Best Local Similarity 100.0%; Score 20; DB 21; Length 20;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dy 1 AGAGATGCCCAACTGTTT 20
 1 AGAGATGCCCAACTGTTT 20

RESULT 2
 ABS77493/C

ID ABS77493 standard; CDNA; 732 BP.
 AC ABS77493;
 XX

DT 12-DEC-2002 (first entry)
 XX

DB Frog embryonic gene sequence Q9925901.
 XX

KW Frog; ss; embryonic development; developmental disorder;
 KW microarray; cell differentiation.
 XX

OS Xenopus laevis.
 XX

PN US002081610-A1.
 XX

PD 27-JUN-2002.
 XX

PF 23-JUL-2001; 2001US-0910943.
 XX

PR 21-JUL-2000; 2000US-219658P.
 XX

PA (UVRQ) UNIV ROCKEFELLER.
 XX

PI Hemmati-Brivanlou A, Altman CR;
 XX

DR WPI; 2002-626534/67.
 XX

PT Nucleic acid array containing Xenopus embryonic nucleic acids is useful
 PT to identify genes involved in embryonic development, to identify
 PT different types of embryonic cells, and to diagnose developmental
 PT disorders
 XX

PS Claim 1; Page 763; 823pp; English.
 XX

CC The invention relates to a nucleic acid array, where each coordinate
 CC contains a single nucleic acid species having one of 770 nucleotide
 CC sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic
 CC product, or its complement or hybridizable fragment of not less than
 CC 20 contiguous nucleotides of one of those sequences. Also included are
 CC detecting differential expression of embryonic genes, comprising:
 CC (a) contacting a nucleic acid array of embryonic genes, comprising:
 CC embryonic but not mature cells with nucleic acids from sample and
 CC control cells; and (b) detecting differential hybridisation of nucleic
 CC acids from the sample cells relative to the control cells; and detecting
 CC defects in development, comprising: (a) contacting nucleic acids from
 CC test cells undergoing development with a nucleic acid array of gene
 CC products known to play a fundamental role in the development process; and
 CC (b) detecting a difference in expression of a fundamental gene in the
 CC sample cells relative to a standard. The invention is useful to identify
 CC genes involved in embryonic development and related processes such as

CC cell differentiation. This would be useful for diagnosing developmental
 CC disorders and for identifying different types of embryonic cells.
 CC The present sequence is one of the 770 Xenopus embryonic CDNA sequences.
 XX

SO Sequence 732 BP; 208 A; 154 C; 200 G; 169 T; 1 other;
 Query Match
 Best Local Similarity 87.0%; Score 17.4; DB 24; Length 732;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dy 2 GAGATGCCCAACTGTTT 20
 101 GAGATGCCCAACTGTTT 83

RESULT 3
 AAX91990/C

ID AAX91990 standard; DNA; 1230025 BP.
 AC AAX91990;
 XX

DT 13-SEP-1999 (first entry)
 XX

DB Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope; ss.
 XX

OS Chlamydia pneumoniae.
 XX

PN W09927105-A2.
 XX

PD 03-JUN-1999.
 XX

PF 20-NOV-1998; 98WO-IB01890.
 XX

PR 04-NOV-1998; 98US-0107078.
 XX

PR 21-NOV-1997; 97FR-0014673.
 XX

PA (GBST) GENSET.
 XX

PI Griffiths R;
 XX

DR WPI; 1999-357842/30.
 XX

PT Genome sequence of Chlamydia pneumoniae
 PT
 XX

PS Claim 1; Page 291-611; 1912pp; English.
 XX

CC The present sequence represents the complete genome of Chlamydia
 CC pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes
 CC respiratory disease such as pneumonia and bronchitis and is thought
 CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,
 CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
 CC encoded by the open reading frames of the C. pneumoniae genome (see
 CC AAY34584-Y35879) can be used in immunogenic compositions as vaccine.
 CC Vectors containing C. pneumoniae nucleotide sequences can also be
 CC used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae.
 XX

SO Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;
 Query Match
 Best Local Similarity 87.0%; Score 17.4; DB 20; Length 1230025;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dy 2 GAGATGCCCAACTGTTT 20
 1034868 GAGATGCCCAACTGTTT 1034850

RESULT 4

AA551323/c
ID AAS51323 standard; DNA; 234 BP.
XX
AC AAS51323;
XX
DT 13-FEB-2002 (first entry)
XX
DE Enterococcus faecalis DNA for cellular proliferation protein #100.
XX
DE Antisense; ds; prokaryotic cellular proliferation gene;
XX
KM antibiotic; antibacterial; drug design.
XX
OS Enterococcus faecalis.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-20727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI; 2001-611495/70.
XX
DR P-PSDB; AAU33464.
XX
PT New polynucleotides for the identification and development of
XX
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 3905; 511P; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
XX
CC prokaryotic cellular proliferation, their use in identifying the
XX
CC genes, their use in the discovery of novel antibiotics, the essential
XX
CC genes themselves and the encoded proteins. The prokaryotes used are
XX
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
XX
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
XX
CC invention is also useful for the identification of potential new targets
XX
CC for antibiotic development. The antisense nucleic acids can also be used
XX
CC to identify proteins used in proliferation, to express these proteins,
XX
CC and to obtain antibodies capable of binding to the expressed proteins.
XX
CC The proteins can be used to screen compounds in rational drug discovery
XX
CC programmes. The antisense nucleic acid sequence is also useful to screen
XX
CC for homologous nucleic acids which are required for cell proliferation in
XX
CC a wide variety of organisms. The present sequence encodes an
XX
CC essential prokaryotic cellular proliferation protein.
XX
CC Note: The sequence data for this patent did not form part
XX
CC of the printed specification, but was obtained in electronic
XX
CC format directly from WIPO at
XX
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 234 BP; 84 A; 33 C; 46 G; 71 T; 0 other;
XX
Query Match 84.0%; Score 16.8; DB 23; Length 234;
XX
Best Local Similarity 90.0%; Pred. No. 63;
XX
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

AA553070/c
ID AAS53070 standard; DNA; 306 BP.
XX
AC AAS53070;
XX
DT 13-FEB-2002 (first entry)
XX
DE Enterococcus faecalis DNA for cellular proliferation protein #498.
XX
DE Antisense; ds; prokaryotic cellular proliferation gene;
XX
KM antibiotic; antibacterial; drug design.
XX
OS Enterococcus faecalis.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-20727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI; 2001-611495/70.
XX
DR P-PSDB; AAU35211.
XX
PT New polynucleotides for the identification and development of
XX
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 6707; 511P; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
XX
CC prokaryotic cellular proliferation, their use in identifying the
XX
CC genes, their use in the discovery of novel antibiotics, the essential
XX
CC genes themselves and the encoded proteins. The prokaryotes used are
XX
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
XX
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
XX
CC invention is also useful for the identification of potential new targets
XX
CC for antibiotic development. The antisense nucleic acids can also be used
XX
CC to identify proteins used in proliferation, to express these proteins,
XX
CC and to obtain antibodies capable of binding to the expressed proteins.
XX
CC The proteins can be used to screen compounds in rational drug discovery
XX
CC programmes. The antisense nucleic acid sequence is also useful to screen
XX
CC for homologous nucleic acids which are required for cell proliferation in
XX
CC a wide variety of organisms. The present sequence encodes an
XX
CC essential prokaryotic cellular proliferation protein.
XX
CC Note: The sequence data for this patent did not form part
XX
CC of the printed specification, but was obtained in electronic
XX
CC format directly from WIPO at
XX
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 306 BP; 117 A; 40 C; 65 G; 84 T; 0 other;
XX
Query Match 84.0%; Score 16.8; DB 23; Length 306;
XX
Best Local Similarity 90.0%; Pred. No. 65;
XX
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 5

RESULT 6

KW cardiovascular disorder; respiratory disorder; neurological disease;
 KW infection; reproductive system disorder; immunosuppressive;
 KW antitubercular; antineoplastic; antiproliferative; cytostatic; cardiac;
 KW vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
 KW virucide; fungicide; ophthalmological; gene; ss.
 XX Homo sapiens.
 XX MO200228877-A1.
 XX 11-APR-2002.
 XX 17-JAN-2001; 2001MO-US01432.
 XX 29-SEP-2000; 2000US-236326P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
 XX Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;
 XX Ni J;
 XX WPI; 2002-340092/37.
 XX P-PSDB; AAO17218.
 XX Isolated nucleic acid molecule encoding a human secreted protein is
 XX used in preventing, treating or ameliorating a medical condition -
 XX Claim 1; Page 478; 538pp; English.
 XX The present invention provides the protein and coding sequences of a
 XX number of human secreted proteins. These can be used in the treatment of
 XX cancer, immune diseases, haemostatic disorders, hyperproliferative
 XX disorders, renal disorders, cardiovascular disorders, respiratory
 XX disorders, neurological diseases, infections and reproductive system
 XX disorders. The present sequence is a coding sequence of the invention.
 XX Sequence 409 BP; 106 A; 62 C; 74 G; 163 T; 4 other;
 XX
 XX Query Match 84.0%; Score 16.8; DB 24; Length 409;
 XX Best Local Similarity 90.0%; Pred. No. 69;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX 1 AGAGATGCCCAACTGTTT 20
 XX |||||
 XX 228 AGAATGCGCAACTGTTT 247
 XX
 XX RESULT 9
 XX AAT03501/C
 XX ID AAT03501 standard; DNA; 661 BP.
 XX
 XX AAT03501;
 XX 12-JUN-1996 (first entry)
 XX
 XX Papilloma virus major capsid protein DNA.
 XX
 XX HP-Virus 49; papilloma virus; major;
 KW capsid protein; plasmid VS40-7; DSM 9135; diagnosis;
 KW skin carcinomas; therapy; vaccination; de.
 XX
 XX Papilloma virus.
 XX
 XX Key Location/Qualifiers
 XX mat_peptide 1..660
 XX /*tag= a
 XX
 XX DE4415743-A1.
 XX
 XX 09-NOV-1995.
 XX
 XX 04-MAY-1994; 94DE-4415743.

XX 04-MAY-1994; 94DE-4415743.
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 XX de Villiers zur Hausen E, Leigh I, Shuman V, zur Hausen H;
 XX
 XX WPI; 1995-383680/50.
 XX P-PSDB; AAR88269.
 XX DNA encoding peptide(s) of papilloma virus major capsid protein -
 XX useful for detecting papilloma virus in skin carcinoma
 XX
 XX Claim 1; Fig 3; 15pp; German.
 XX
 XX AAT03501 is 78.5% homologous to HP-Virus 49, encodes AAR88269 a
 XX peptide of papilloma virus (PV) major capsid protein and is
 XX expressed by the plasmid VS40-7 (DSM 9135). The DNA is useful
 XX in diagnosis, esp. to detect PV in skin carcinomas. Major capsid
 XX proteins encoded by a PV genome contg. the DNA are useful for
 XX therapy and/or vaccination.
 XX
 XX Sequence 661 BP; 195 A; 118 C; 155 G; 193 T; 0 other;
 XX
 XX Query Match 84.0%; Score 16.8; DB 16; Length 661;
 XX Best Local Similarity 90.0%; Pred. No. 74;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX 1 AGAGATGCCCAACTGTTT 20
 XX |||||
 XX 567 AGAGTTCGCCAAGTGT 548
 XX
 XX RESULT 10
 XX AAL45853
 XX ID AAL45853 standard; cDNA; 1569 BP.
 XX
 XX AAL45853;
 XX 08-JUL-2002 (first entry)
 XX
 XX Human secreted protein coding sequence SEQ ID NO: 58.
 XX
 XX Human; secreted protein; gene therapy; cancer; immune disease;
 KW haemostatic disorder; hyperproliferative disorder; renal disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disease;
 KW infection; reproductive system disorder; immunosuppressive;
 KW antitubercular; antineoplastic; antiproliferative; cytostatic; cardiac;
 KW vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
 KW virucide; fungicide; ophthalmological; gene; ss.
 XX
 XX Homo sapiens.
 XX
 XX MO200228877-A1.
 XX 11-APR-2002.
 XX
 XX 17-JAN-2001; 2001MO-US01432.
 XX 29-SEP-2000; 2000US-236326P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
 XX Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;
 XX Ni J;
 XX WPI; 2002-340092/37.
 XX P-PSDB; AAO17217.
 XX Isolated nucleic acid molecule encoding a human secreted protein is
 XX used in preventing, treating or ameliorating a medical condition -

PS Claim 1; Page 478; 538pp; English.

CC The present invention provides the protein and coding sequences of a
 CC number of human secreted proteins. These can be used in the treatment of
 CC cancer, immune diseases, haemostatic disorders, hyperproliferative
 CC disorders, renal disorders, cardiovascular disorders, respiratory
 CC disorders, neurological diseases, infections and reproductive system
 CC disorders. The present sequence is a coding sequence of the invention.
 XX

SO Sequence 1569 BP; 397 A; 235 C; 236 G; 691 T; 10 other;

Query Match 84.0%; Score 16.8; DB 24; Length 1569;
 Best Local Similarity 90.0%; Pred. No. 86;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAACTGTTT 20
 228 AGAATGCGCAACTGTTT 247

RESULT 11

AAV39297 standard; cDNA; 2008 BP.

AC AAV39297;

DT 16-SEP-1998 (first entry)

DE Human RAD54 nucleic acid sequence comprising exons 5-8.

XX Human; RAD54; cancer; xeroderma pigmentosum; Bloom syndrome;
 KW Werner's syndrome; ATR-X; diagnosis; detection; SNF2 superfamily;
 KM X-linked mental retardation with alpha-thalassemia syndrome; tumour;
 KW gene therapy; ss.

XX Homo sapiens.

OS Homo sapiens.

PN EP844305-A2.

PD 27-MAY-1998.

PF 10-NOV-1997; 97EP-0308998.

PR 13-NOV-1996; 96US-0030676.

XX (SMK) SMITHKLINE BEECHAM CORP.

PA (UIVE-) UNIV JEFFERSON THOMAS.

PI Croce CM, Fishel RA, Rasio D, Robbins DJ;

DR WPI; 1998-274189/25.

XX Human hRAD54 DNA and polypeptide - and agonists, antibodies,
 XX antagonists, etc.

PT Human hRAD54 DNA and polypeptide - and agonists, antibodies,
 XX antagonists, etc.

PS Claim 1; Page 27-28; 64pp; English.

CC The present sequence represents a specifically claimed partial nucleic
 CC acid sequence encoding human RAD54 (hRAD54). A method for analysing a
 CC sample for mutation of DNA encoding hRAD54 has been developed using a
 CC DNA sequence of at least 15 and no more than 30 consecutive bases of
 CC the DNA sequence encoding hRAD54. hRAD54 is a gene thought to be present
 CC in tumours that display allelic imbalance at 1p32, the chromosomal band
 CC identified as one of four minimal regions of chromosome 1 deletion in
 CC breast carcinomas. hRAD54 is useful for production of proteins, inter
 CC alia, that have been identified as novel hRAD54 by homology between the
 CC amino acid sequence given in AAW62186 and known amino acid sequences
 CC such as yeast RAD54. hRAD54 proteins are used in the treatment of
 CC cancer, including xeroderma pigmentosum and Bloom syndrome, Werner's
 CC syndrome and X-linked mental retardation with alpha-thalassemia
 CC syndrome and breast cancer. hRAD54 polynucleotides are also useful for
 CC detecting complementary nucleotides for use as a diagnostic agent,
 CC especially useful for diagnosis of disease or susceptibility to diseases.

CC hRAD54 polynucleotide, proteins, agonists and antagonists which are
 CC proteins are useful in gene therapy.

SO Sequence 2008 BP; 489 A; 419 C; 570 G; 526 T; 4 other;

Query Match 84.0%; Score 16.8; DB 19; Length 2008;
 Best Local Similarity 90.0%; Pred. No. 90;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAACTGTTT 20
 139 AGAGATGCCCAACTGATT 158

RESULT 12

AA145822 standard; cDNA; 2185 BP.

AC AA145822;

DT 08-JUL-2002 (first entry)

DE Human secreted protein coding sequence SEQ ID NO: 27.

XX Human; secreted protein; gene therapy; cancer; immune disease;
 KW haemostatic disorder; hyperproliferative disorder; renal disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disease;
 KW infection; reproductive system disorder; immunosuppressive;
 KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
 KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
 KW virucide; fungicide; ophthalmological; gene; ss.

XX Homo sapiens.

PN WO200228877-A1.

PD 11-APR-2002.

PF 17-JAN-2001; 2001WO-US01432.

PR 29-SEP-2000; 2000US-236326P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Komatsu's GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

PI Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;

PI Ni J;

DR WPI; 2002-340092/37.

DR P-PsDB; AAO17186.

PT Isolated nucleic acid molecule encoding a human secreted protein is
 XX used in preventing, treating or ameliorating a medical condition

PS Claim 1; Page 458-459; 538pp; English.

CC The present invention provides the protein and coding sequences of a
 CC number of human secreted proteins. These can be used in the treatment of
 CC cancer, immune diseases, haemostatic disorders, hyperproliferative
 CC disorders, renal disorders, cardiovascular disorders, respiratory
 CC disorders, neurological diseases, infections and reproductive system
 CC disorders. The present sequence is a coding sequence of the invention.
 XX

SO Sequence 2185 BP; 613 A; 302 C; 320 G; 950 T; 0 other;

Query Match 84.0%; Score 16.8; DB 24; Length 2185;
 Best Local Similarity 90.0%; Pred. No. 91;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAACTGTTT 20
 228 AGAATGCGCAACTGTTT 247

RESULT 13
AAH51750/c
ID AAH51750 standard; DNA; 3001 BP.

AAH51750;
29-AUG-2001 (first entry)

Chromosome 13q31-q33 biallelic marker containing amplicon SEQ ID 162.

sbgl; g34665; sbg2; g35017; g35018; Chromosome 13q31-q33; haplotype; bi-allelic marker; polymorphism; schizophrenia; bipolar disorder; ds.

Homo sapiens.

MO200058510-A2.

05-OCT-2000.

30-MAR-2000; 2000MO-IB00435.

30-MAR-1999; 99US-0126903.

30-APR-1999; 99US-0131971.

30-APR-1999; 99US-0132065.

14-JUL-1999; 99US-0143928.

27-JUL-1999; 99US-0145915.

29-JUL-1999; 99US-0146452.

29-JUL-1999; 99US-0146453.

28-OCT-1999; 99US-0162288.

(GIST) GENSET.

Cohen D, Blumenfeld M, Chumakov I, Bouguetelret L, Bihain B;

Essioux L;

WPI; 2000-619082/59.

Polynucleotides comprising sequences from sbgl and g35018 bi-allelic

markers are used for genotyping and detecting schizophrenia or bipolar

disorder and predisposition to these disorders -

Claim 2; Page 643-644; 737PD; English.

AAH51601 represents a human genomic nucleotide sequence comprising sbgl,

g34665, sbg2, g35017 and g35018 nucleic acid sequences located on the

human chromosome 13q31-q33 locus. The nucleotide sequences contain

bi-allelic markers and polymorphisms. Sequences AAH51602 - AAH51626 and

AAH51627 - AAH51631 represent cDNA human sbgl cDNA sequences and protein

products. AAH51632 - AAH51636 represent cDNA human sbgl cDNA sequences and protein

products. AAH51637 - AAH51641 represent cDNA human sbgl cDNA sequences and protein

products. AAH51642 - AAH51646 represent cDNA human sbgl cDNA sequences and protein

products. AAH51647 - AAH51651 represent cDNA human sbgl cDNA sequences and protein

products. AAH51652 - AAH51656 represent cDNA human sbgl cDNA sequences and protein

products. AAH51657 - AAH51661 represent cDNA human sbgl cDNA sequences and protein

products. AAH51662 - AAH51666 represent cDNA human sbgl cDNA sequences and protein

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20
db 427 AGAGCTGCCCAACTGTTT 408

RESULT 14
AAH03245/c
ID AAH03245 standard; DNA; 9821 BP.

AAH03245;
21-NOV-2001 (first entry)

Human reproductive system related antigen DNA SEQ ID NO: 5933.

Human reproductive system related antigen; reproductive system disorder;

Human; gene therapy; ds.

Homo sapiens.

MO200155320-A2.

02-AUG-2001.

17-JAN-2001; 2001MO-US01339.

31-JUN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0215135.

30-JUN-2000; 2000US-0216647.

07-JUL-2000; 2000US-0216880.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225266.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225457.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0225759.

```

PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246538.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.

```

```

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251859.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR XX
PR PA (HUMA-) HUMAN GENOME SCI INC.
PR XX
PR PI Rosen CA, Barash SC, Ruben SM;
PR DR WPI; 2001-465570/50.
PR XX
PR PT Isolated nucleic acid molecule encoding a reproductive system antigen
PR PT is used in preventing, treating or ameliorating a medical condition -
PR XX
PR PS Disclosure, SEQ ID NO 5933; 1297bp + Sequence Listing; English.
PR XX
PR CC The present invention provides the protein and coding sequences of a
PR CC number of human reproductive system related antigens. These can be used
PR CC in the prevention and treatment of reproductive system disorders,
PR CC including cancer. The present sequence is a genomic sequence encoding a
PR CC protein of the invention.
PR XX
PR SQ Sequence 9821 BP; 3056 A; 1735 C; 1689 G; 3341 T; 0 other;

```

```

Query Match 84.0%; Score 16.8; DB 22; Length 9821;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGAGATGCCCAACTGCTTT 20
DB 2800 AGAGTCTCCCAACTGCTTT 2781

```

```

RESULT 15
AAL03246/c
ID AAL03246 standard; DNA; 9824 BP.
XX
AC AAL03246;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 5934.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.

```


CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX

XX

Sequence 9824 BP; 3057 A; 1735 C; 1689 G; 3343 T; 0 other;

Query Match
Best 1.0001

Best Local Match	84.0%;	Score 16.8;	DB 22;	Length 9824;
Similarity	90.0%;			
Matches	18;	Pred. No. 1.2e+02;		

```

      Local Similarity  90.0%;  Pred. No. 1.2e+02;
Matches  18;  Conservative  0;  Mismatches  2

```

1 AGAGATGCCCAACTGTTT 20

Db 2800 AGAGCTGCCCAACTGCTTT 2781

Search completed: August 1, 2003, 13:55:17
Job time : 9.69868 secs

Sun Aug 3 09:03:50 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds
(without alignments)
7955.924 Million cell updates/sec

Title: US-09-834-291-12
Perfect score: 20
Sequence: 1 agagatgcccaactgtttc 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	20	100.0	20 9 US-09-834-291-12	Sequence 18, Appl
2	20	100.0	40 9 US-09-834-291-18	Sequence 18, Appl
3	20	100.0	40 9 US-09-834-291-26	Sequence 26, Appl
4	20	100.0	2827 9 US-09-834-291-4	Sequence 4, Appl
5	20	100.0	3212 9 US-09-834-291-1	Sequence 1, Appl
6	17.4	87.0	732 9 US-09-834-291-719	Sequence 719, Appl
7	16.8	84.0	20 9 US-09-834-291-13	Sequence 13, Appl
8	16.8	84.0	20 9 US-09-815-242-1905	Sequence 3905, Ap
9	16.8	84.0	388 10 US-09-974-300-1452	Sequence 6707, Ap
10	16.8	84.0	388 10 US-09-974-300-1452	Sequence 105156, Ap
11	16.8	84.0	650 13 US-10-027-632-205156	Sequence 5933, Ap
12	16.8	84.0	9821 11 US-09-764-891-5933	Sequence 5933, Ap
13	16.8	84.0	9824 11 US-09-764-891-5934	Sequence 89, Appl
14	16.8	84.0	20072 10 US-09-070-927A-89	Sequence 1867, Ap
15	15.8	79.0	472 11 US-09-918-995-1867	Sequence 254882, Ap
16	15.8	79.0	893 13 US-10-027-632-254882	Sequence 254882, Ap

us-09-834-291-12.rnpb

17	15.8	79.0	1195	9	US-09-759-143-878
18	15.8	79.0	1195	9	US-09-780-669-878
19	15.8	79.0	1195	9	US-09-822-827-878
20	15.8	79.0	1195	10	US-09-895-793-878
21	15.8	79.0	1195	10	US-09-895-814-878
22	15.8	79.0	1195	13	US-10-012-896-878
23	15.8	79.0	1195	14	US-10-011-095-1
24	15.8	79.0	1195	14	US-10-010-667A-1
25	15.8	79.0	1195	14	US-10-205-823-396
26	15.8	79.0	2104	9	US-09-842-552-99
27	15.8	79.0	3627	14	US-10-011-095-6
28	15.8	79.0	32172	11	US-10-010-667A-6
29	15.8	79.0	2140405	13	US-09-764-891-6161
30	15.8	79.0	368	10	US-10-027-632-76212
31	15.4	77.0	646	13	US-09-783-590-1742
32	15.4	77.0	773	13	US-10-027-632-142565
33	15.4	77.0	2690	13	US-10-027-632-137412
34	15.4	77.0	110096	10	US-10-027-632-111780
35	15.4	77.0	123526	11	US-09-880-107-1542
36	15.2	76.0	136	11	US-09-910-185-11
37	15.2	76.0	136	11	US-09-747-377-316
38	15.2	76.0	327	11	US-10-105-613-316
39	15.2	76.0	353	11	US-09-803-719-899
40	15.2	76.0	369	13	US-09-803-719-899
41	15.2	76.0	377	10	US-10-027-632-19177
42	15.2	76.0	428	13	US-09-867-701-2517
43	15.2	76.0	448	13	US-10-027-632-60626
44	15.2	76.0	455	13	US-10-027-632-72196
45	15.2	76.0	455	13	US-10-027-632-188985

ALIGNMENTS

RESULT 1
US-09-834-291-12
Sequence 12, Application US/09834291
GENERAL INFORMATION:
PATENT NO. US2002042064A1
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
CURRENT FILING NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-12
Query Match
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0;
Matches 20; Conservative 0; Mismatches 0;
DB 1 AGAGATGCCCAACTGTTT 20
1 AGAGATGCCCAACTGTTT 20
RESULT 2
US-09-834-291-18
Sequence 18, Application US/09834291
GENERAL INFORMATION:
PATENT NO. US2002042064A1
APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina
 TITLE OF INVENTION: P53 Binding Areas
 FILE REFERENCE: 4121-122
 CURRENT APPLICATION NUMBER: US/09/834,291
 PRIOR FILING DATE: 2001-08-21
 PRIOR APPLICATION NUMBER: PCT/DE99/03343
 PRIOR FILING DATE: 1999-10-18
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 18
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Homo Sapiens
 US-09-834-291-18

Query Match
 Best Local Similarity 100.0%; Score 20; DB 9; Length 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
 DB 11 AGAGATGCCCAACTGTTT 30

RESULT 3
 US-09-834-291-26
 Sequence 26, Application US/09834291
 Patent No. US20020042064A1
 GENERAL INFORMATION:
 APPLICANT: Krammer, Peter
 APPLICANT: Muller-Schilling, Martina
 TITLE OF INVENTION: P53 Binding Areas
 FILE REFERENCE: 4121-122
 CURRENT APPLICATION NUMBER: US/09/834,291
 PRIOR FILING DATE: 2001-08-21
 PRIOR APPLICATION NUMBER: PCT/DE99/03343
 PRIOR FILING DATE: 1999-10-18
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 26
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Homo Sapiens
 US-09-834-291-26

Query Match
 Best Local Similarity 100.0%; Score 20; DB 9; Length 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
 DB 11 AGAGATGCCCAACTGTTT 30

RESULT 4
 US-09-834-291-4
 Sequence 4, Application US/09834291
 Patent No. US20020042064A1
 GENERAL INFORMATION:
 APPLICANT: Krammer, Peter
 APPLICANT: Muller-Schilling, Martina
 TITLE OF INVENTION: P53 Binding Areas
 FILE REFERENCE: 4121-122
 CURRENT APPLICATION NUMBER: US/09/834,291
 PRIOR FILING DATE: 2001-08-21
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Homo Sapiens
 US-09-834-291-4

PRIOR FILING DATE: 1999-10-18
 PRIOR APPLICATION NUMBER: DE 198 47 779.1
 PRIOR FILING DATE: 1998-10-16
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 2827
 TYPE: DNA
 ORGANISM: Homo Sapiens
 US-09-834-291-4

Query Match
 Best Local Similarity 100.0%; Score 20; DB 9; Length 2827;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
 DB 44 AGAGATGCCCAACTGTTT 63

RESULT 5
 US-09-834-291-1
 Sequence 1, Application US/09834291
 Patent No. US20020042064A1
 GENERAL INFORMATION:
 APPLICANT: Krammer, Peter
 APPLICANT: Muller-Schilling, Martina
 TITLE OF INVENTION: P53 Binding Areas
 FILE REFERENCE: 4121-122
 CURRENT APPLICATION NUMBER: US/09/834,291
 PRIOR FILING DATE: 2001-08-21
 PRIOR APPLICATION NUMBER: PCT/DE99/03343
 PRIOR FILING DATE: 1999-10-18
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 3212
 TYPE: DNA
 ORGANISM: Homo Sapiens
 US-09-834-291-1

Query Match
 Best Local Similarity 100.0%; Score 20; DB 9; Length 3212;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
 DB 44 AGAGATGCCCAACTGTTT 63

RESULT 6
 US-09-910-943-719/c
 Sequence 719, Application US/09910943
 Patent No. US20020081610A1
 GENERAL INFORMATION:
 APPLICANT: Hemmati-Brivanlou, Ali
 TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
 FILE REFERENCE: 7529/1G148U31
 CURRENT APPLICATION NUMBER: US/09/910,943
 PRIOR FILING DATE: 2001-07-23
 NUMBER OF SEQ ID NOS: 742
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 719
 LENGTH: 732
 TYPE: DNA
 ORGANISM: Xenopus laevis
 FEATURE: NAME/KEY: misc_feature
 LOCATION: (1)..(732)

OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-719

Query Match 87.0%; Score 17.4; DB 9; Length 732;
Best Local Similarity 94.7%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20
DB 101 GAGATGCCCAACTGTTCT 83

RESULT 7
US-09-834-291-13
Sequence 13, Application US/09834291
Patent No. US20020042064A1

GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-13

Query Match 84.0%; Score 16.8; DB 9; Length 20;
Best Local Similarity 90.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
DB 1 AGAGATGCCCAAAATGTTT 20

RESULT 8
US-09-815-242-3905/c

Sequence 3905, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3905
LENGTH: 234
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-815-242-3905

Query Match 84.0%; Score 16.8; DB 9; Length 234;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
DB 87 AGAATGCACCAACTGTTT 68

RESULT 9
US-09-815-242-6707/c

Sequence 6707, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6707
LENGTH: 306
TYPE: DNA
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(306)
US-09-815-242-6707

Query Match 84.0%; Score 16.8; DB 9; Length 306;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
DB 105 AGAATGCACCAACTGTTT 86

```
RESULT 10
US-09-974-300-1452
; Sequence 1452, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10095.500-US
; CURRENT APPLICATION NUMBER: US/09/974.300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1452
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1452
```

```
Query Match
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 368;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AGAGATGCCCAAACTGTTT 20
Db 342 AGAGATGCCCAAACTGTTT 361
```

```
RESULT 11
US-10-027-632-205156
; Sequence 205156, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 07/60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205156
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-205156
```

```
Query Match
Best Local Similarity 84.0%; Score 16.8; DB 13; Length 650;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AGAGATGCCCAAACTGTTT 20
Db 70 AGAGATGCCCAAACTGTTT 89
```

```
RESULT 12
US-09-764-891-5933/c
; Sequence 5933, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5933
; LENGTH: 9821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5933
```

```
Query Match
Best Local Similarity 84.0%; Score 16.8; DB 11; Length 9821;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AGAGATGCCCAAACTGTTT 20
Db 2800 AGAGATGCCCAAACTGTTT 2781
```

```
RESULT 13
US-09-764-891-5934/c
; Sequence 5934, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5934
; LENGTH: 9824
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5934
```

```
Query Match
Best Local Similarity 84.0%; Score 16.8; DB 11; Length 9824;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AGAGATGCCCAAACTGTTT 20
Db 2800 AGAGATGCCCAAACTGTTT 2781
```

```
RESULT 14
US-09-070-927A-89
; Sequence 89, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
```

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A

FILING DATE: 04-May-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:

LENGTH: 20072 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 89:

US-09-070-927A-89

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 10;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTGTTT 20

DB 7383 AGAATGCACAACTGTGTTT 7402

RESULT 15

US-09-918-995-1867

Sequence 1867, Application US/0918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: HySeq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1867

LENGTH: 472

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(472)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-1867

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 11;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTGTTT 20

DB 73 GAGATGCCCTACTGTGTTT 91

Search completed: August 1, 2003, 13:36:57

Job time : 6.18607 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.955 Million cell updates/sec

Title: US-09-834-291-12

Sequence: 1 agagatgcccaactgtttc 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
1	17.4	87.0 1230025	4 US-09-198-452A-1	Sequence 1, Appli
2	16.8	84.0 661	3 US-08-578-634C-3	Sequence 3, Appli
3	16.8	84.0 661	4 US-09-430-010-3	Sequence 3, Appli
4	16.8	84.0 3001	4 US-09-539-333D-162	Sequence 162, App
5	16.8	84.0 43795	3 US-08-742-185-101	Sequence 101, App
6	15.8	79.0 1195	4 US-09-323-873A-1	Sequence 1, Appli
7	15.8	79.0 3627	4 US-07-903-047-7	Sequence 7, Appli
8	15.4	77.0 1644	4 US-09-111-752-13	Sequence 13, Appli
9	15.4	77.0 1644	3 US-09-380-061B-15	Sequence 15, Appli
10	15.4	77.0 1644	3 US-08-487-183A-13	Sequence 13, Appli
11	15.4	77.0 1704	4 US-08-460-934-5	Sequence 5, Appli
12	15.4	77.0 1704	2 US-08-782-118-5	Sequence 8, Appli
13	15.4	77.0 1908	1 US-08-460-934-8	Sequence 8, Appli
14	15.4	77.0 1908	2 US-08-782-118-8	Sequence 8, Appli
15	15.4	77.0 1664976	4 US-08-916-421B-1	Sequence 1, Appli
16	15.4	77.0 1664976	4 US-08-916-421B-1	Sequence 1, Appli
17	15.4	77.0 1664976	4 US-08-916-421B-1	Sequence 1, Appli
18	15.2	76.0 450	4 US-09-134-001C-2038	Sequence 2038, Ap
19	15.2	76.0 668	3 US-08-578-634C-2	Sequence 2, Appli
20	15.2	76.0 668	3 US-09-430-010-2	Sequence 2, Appli
21	15.2	76.0 927	4 US-09-221-017B-753	Sequence 753, App
22	15.2	76.0 1185	4 US-09-328-352-1283	Sequence 1283, App
23	15.2	76.0 3970	4 US-09-499-203-14	Sequence 14, Appli
24	15.2	76.0 6204	4 US-09-499-203-20	Sequence 20, Appli
25	15.2	76.0 6792	4 US-09-374-454-20	Sequence 20, Appli
26	15.2	76.0 7387	4 US-09-499-203-17	Sequence 17, Appli
27	15.2	76.0 9321	4 US-09-499-203-1	Sequence 1, Appli

28	15.2	76.0 99500	4 US-09-798-096-10	Sequence 10, Appli
29	15.2	76.0 148567	4 US-09-801-876B-3	Sequence 3, Appli
30	15.2	76.0 319608	4 US-09-539-333D-1	Sequence 1, Appli
31	15.2	76.0 319608	4 US-09-679-409-1	Sequence 1, Appli
32	14.8	74.0 432	4 US-09-641-638-435	Sequence 435, App
33	14.8	74.0 495	3 US-09-109-204-20	Sequence 20, Appli
34	14.8	74.0 495	4 US-09-490-032-20	Sequence 19, Appli
35	14.8	74.0 605	3 US-09-109-204-19	Sequence 19, Appli
36	14.8	74.0 605	4 US-09-490-032-19	Sequence 19, Appli
37	14.8	74.0 660	4 US-08-961-527-246	Sequence 246, App
38	14.8	74.0 1001	4 US-09-671-317-240	Sequence 240, App
39	14.8	74.0 2125	3 US-09-109-204-5	Sequence 5, Appli
40	14.8	74.0 2125	4 US-09-490-032-5	Sequence 5, Appli
41	14.8	74.0 3794	4 US-09-608-821-1	Sequence 1, Appli
42	14.8	74.0 3804	4 US-08-842-382-1	Sequence 1, Appli
43	14.8	74.0 4712	4 US-09-221-017B-577	Sequence 577, App
44	14.8	74.0 26654	4 US-09-564-805-28	Sequence 28, Appli
45	14.4	72.0 832	1 US-08-473-496-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-198-452A-1/c
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffeis, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (150001)..(165000)
```

[illegible]

Sun Aug 3 09:03:49 2003

us-09-834-291-12.rnt

LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

Query Match 87.0%; Score 17.4; DB 4; length 1230025;
Best Local Similarity 94.7%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20
Db 1034868 GAGATGCCCAACTGTTT 1034850

RESULT 2
US-08-578-634C-3/C
Sequence 3, Application US/08578634C
Patent No. 6025163

GENERAL INFORMATION:
APPLICANT: Vladimir Shamanin
APPLICANT: Echel Michele De Villiers
APPLICANT: Zur Hausen
APPLICANT: Irene Leigh
APPLICANT: Harald Zur Hausen
TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA
TITLE OF INVENTION: VIRUS MAIN CAPSIDE PROTEIN AND USE THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/578,634C
FILING DATE: 26-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B.
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8484-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)493-4935
TELEFAX: (650)493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..660

US-08-578-634C-3
Query Match 84.0%; Score 16.8; DB 3; length 661;
Best Local Similarity 90.0%; Pred. No. 9.2;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20

Db 567 AGAGTCCCAAGTGT 548

RESULT 3
US-09-430-010-3/C
Sequence 3, Application US/09430010
Patent No. 6395512

GENERAL INFORMATION:
APPLICANT: Vladimir Shamanin
APPLICANT: Echel Michele De Villiers
APPLICANT: Zur Hausen
APPLICANT: Irene Leigh
APPLICANT: Harald Zur Hausen
TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA
TITLE OF INVENTION: VIRUS MAIN CAPSIDE PROTEIN AND USE THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,010
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/578,634
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B.
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8484-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)493-4935
TELEFAX: (650)493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..660

US-09-430-010-3
Query Match 84.0%; Score 16.8; DB 4; length 661;
Best Local Similarity 90.0%; Pred. No. 9.2;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20
Db 567 AGAGTCCCAAGTGT 548

RESULT 4
US-09-539-333D-162/C
Sequence 162, Application US/09539333D
Patent No. 6476208

GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya

APPLICANT: Bougueleret, Lydie
APPLICANT: Bihain, Bernard
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLERIC MARKERS
FILE REFERENCE: GENSET.047AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US 60/126,903
PRIOR FILING DATE: 1999-03-30
PRIOR APPLICATION NUMBER: US 60/131,971
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: US 60/132,065
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: US 60/143,928
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: US 60/145,915
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: US 60/146,453
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: US 60/146,452
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: US 60/162,288
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: US 09/416,384
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patent.pm
SEQ ID NO 162
LENGTH: 3001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 1501
OTHER INFORMATION: 8-94-252 : polymorphic base A or G
FEATURE:
NAME/KEY: misc.binding
LOCATION: 1482..1500
OTHER INFORMATION: 8-94-252.misl
FEATURE:
NAME/KEY: misc.binding
LOCATION: 1502..1521
OTHER INFORMATION: 8-94-252.mls2, complement
FEATURE:
NAME/KEY: primer.bind
LOCATION: 1250..1267
OTHER INFORMATION: upstream amplification primer
FEATURE:
NAME/KEY: primer.bind
LOCATION: 1651..1669
OTHER INFORMATION: downstream amplification primer, complement
FEATURE:
NAME/KEY: misc.binding
LOCATION: 1489..1513
OTHER INFORMATION: 8-94-252 probe
US-09-539-333D-162

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 4; Length 3001;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAACTGTTT 20
427 AGAGCTGCCCAACTGCTTT 408

RESULT 5
US-08-742-185-101
Sequence 101 Application US/08742185
GENERAL INFORMATION:
APPLICANT: Page, David C.
APPLICANT: Reijo, Renee

APPLICANT: Saxena, Richa
APPLICANT: Hawkins, Trevor
APPLICANT: Reeve, Mary Pat
TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,185
FILING DATE: 30-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/690,734
FILING DATE: 31-JUL-1996
APPLICATION DATA:
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 33,227
REFERENCE/DOCKET NUMBER: WH194-07A2
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 43795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-742-185-101

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 3; Length 43795;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAACTGTTT 20
22516 AGAGATTGCCCAACTGTTT 22535

RESULT 6
US-09-323-873A-1/c
Sequence 1 Application US/09323873A
Patent No. 6329503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Arthur B. Raitano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 129.16US02
CURRENT APPLICATION NUMBER: US/09/323,873A
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1195
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-323-873A-1

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 4; Length 1195;
Pred. No. 35;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20
DB 566 GAGATGCCCAACTGCTTT 548

RESULT 7
US-09-323-873A-6/c
Sequence 6, Application US/0932873A
Patent No. 632503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahan Beong
APPLICANT: Arthur B. Kaitano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
FILE REFERENCE: 129.16USU2
CURRENT APPLICATION NUMBER: US/09/323, 873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087, 520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091, 183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 3627
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-323-873A-6

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 4; Length 3627;
Pred. No. 42;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20
DB 596 GAGATGCCCAACTGCTTT 578

RESULT 8
US-07-903-047-7/c
Sequence 7, Application US/07903047
Patent No. 5228285
GENERAL INFORMATION:
APPLICANT: Kajiyama, Naoki
APPLICANT: Nakano, Eiichi
TITLE OF INVENTION: Thermostable Luciferase Gene Of Firefly
TITLE OF INVENTION: Thermostable Luciferase Gene Of Firefly, No. 5229285el Recombi
TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
TITLE OF INVENTION: Luciferase Of Firefly
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,047
FILING DATE: 19920623
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cDNA to mRNA
US-07-903-047-7

Query Match
Best Local Similarity 77.0%; Score 15.4; DB 1; Length 1644;
Pred. No. 60;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGTGGCCCAACTGT 17
DB 369 AGAGTGGCTTAACGT 353

RESULT 9
US-09-111-752-13/c
Sequence 13, Application US/09111752
Patent No. 6074859
GENERAL INFORMATION:
APPLICANT: HIROKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MURAKAMI, SEIJI
TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111, 752
FILING DATE: 08-JUL-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-0009-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Luciola lateralis
US-09-111-752-13

Query Match 77.0%; Score 15.4; DB 3; Length 1644;
Best Local Similarity 94.1%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17
Db 369 AGAGATGCTTAAGTGT 353

RESULT 10
US-09-380-061B-15/C
Sequence 15, Application US/09380061B
Patent No. 6265177
GENERAL INFORMATION:
APPLICANT: SOUIRELL, DAVID JAMES
LOWE, CHRISTOPHER ROBIN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SB98/01026
FILING DATE: 7-APR-1998
APPLICATION NUMBER: GB 9707468.8
FILING DATE: 11-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B. J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 124-725
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1644
SEQUENCE DESCRIPTION: SEQ ID NO: 15;
US-09-380-061B-15

Query Match 77.0%; Score 15.4; DB 3; Length 1644;
Best Local Similarity 94.1%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17

Db 369 AGAGATGCTTAAGTGT 353

RESULT 11
US-08-487-183A-13/C
Sequence 13, Application US/08487183A
Patent No. 6387675
GENERAL INFORMATION:
APPLICANT: WOOD, Keith V.
APPLICANT: GRUBER, Monika G.
TITLE OF INVENTION: MUTANT LUCIFERASES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: P.O. Box 1497
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-1497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/467,773
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/177,081
FILING DATE: 03-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Scanlon, William J.
REGISTRATION NUMBER: 31,136
REFERENCE/DOCKET NUMBER: 19017/166
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608)258-4258
TELEFAX: (608)258-5035
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1644
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..1644
US-08-487-183A-13

Query Match 77.0%; Score 15.4; DB 4; Length 1644;
Best Local Similarity 94.1%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17
Db 369 AGAGATGCTTAAGTGT 353

RESULT 12
US-08-460-934-5/C
Sequence 5, Application US/08460934
Patent No. 581465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU

APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1704
NAME/KEY: misc feature
LOCATION: 1..1704
OTHER INFORMATION: /note="Nucleotide sequence of the
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombinant
OTHER INFORMATION: plasmid pHLf203 DNA"
US-08-460-934-5
Query Match 77.0%; Score 15.4; DB 1; Length 1704;
Best Local Similarity 94.1%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGAGATGCCCAACTGT 17
Db 429 AGAGATGCCCTAACTGT 413
RESULT 13
US-08-782-118-5/c
; Sequence 5, Application US/08782118
; Patent No. 5843746
; GENERAL INFORMATION:
; APPLICANT: TATSUMI, HIROKI
; APPLICANT: FUKUDA, SATOSHI

APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1704
NAME/KEY: misc feature
LOCATION: 1..1704
OTHER INFORMATION: /note="Nucleotide sequence of the
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombinant
OTHER INFORMATION: plasmid pHLf203 DNA"
US-08-782-118-5
Query Match 77.0%; Score 15.4; DB 2; Length 1704;
Best Local Similarity 94.1%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGAGATGCCCAACTGT 17
Db 429 AGAGATGCCCTAACTGT 413
RESULT 14
US-08-460-934-8/c
; Sequence 8, Application US/08460934
; Patent No. 5814465

GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1908 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1908
OTHER INFORMATION: /note= "The nucleotide sequence of
the biotinylated firefly luciferase gene contained in
OTHER INFORMATION: recombinant plasmid pHLf248 DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1908
US-08-460-934-8

Query Match 77.0%; Score 15.4; DB 1; Length 1908;
Best Local Similarity 94.1%; Pred. No. 62;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17
|||||
Db 369 AGAGATGCCCAACTGT 353

RESULT 15
US-08-782-118-8/c
; Sequence 8, Application US/08782118

Patent No. 5843746
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1908 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1908
OTHER INFORMATION: /note= "The nucleotide sequence of
the biotinylated firefly luciferase gene contained in
OTHER INFORMATION: recombinant plasmid pHLf248 DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1908
US-08-782-118-8

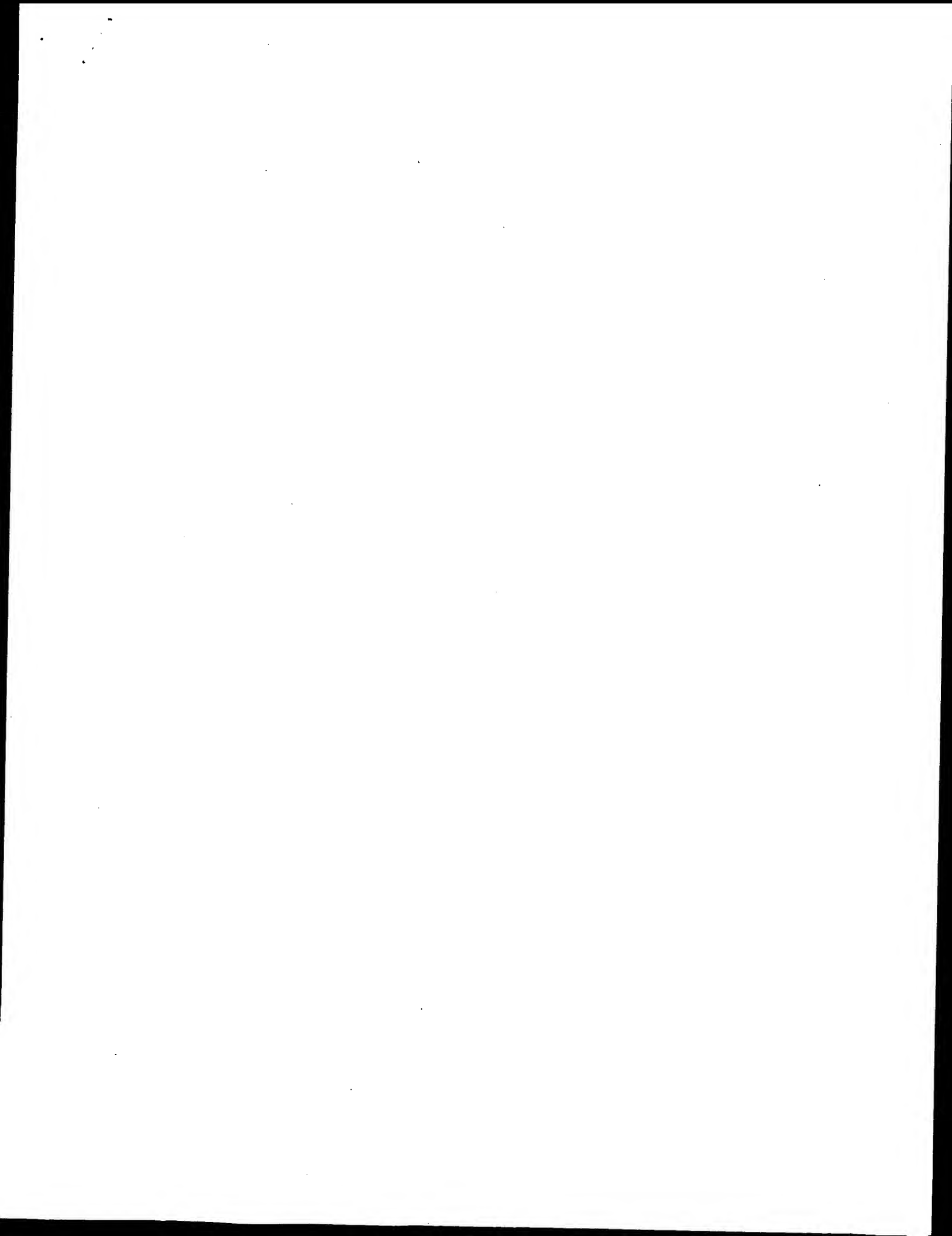
Query Match 77.0%; Score 15.4; DB 2; Length 1908;
Best Local Similarity 94.1%; Pred. No. 62;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17
|||||
Db 369 AGAGATGCCCAACTGT 353

Sun Aug 3 09:03:49 2003

Search completed: August 1, 2003, 08:37:13
Job time : 5.65066 secs

us-09-834-291-12.rml



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds
(without alignments)
10468.541 Million cell updates/sec

Title: US-09-834-291-14
Perfect score: 20
Sequence: 1 aatgttcttaagcttttc 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

GenBank1:
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pa:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vl:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pa:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rod:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Length	DB	ID	Description
1	20	100.0	20	6	AX026102	AX026102 Sequence
2	20	100.0	40	6	AX026108	AX026108 Sequence
3	20	100.0	40	6	AX026116	AX026116 Sequence
4	20	100.0	2344	6	AX026155FR	AX026155FR Sequence
5	20	100.0	2827	6	AX026092	AX026092 Sequence
6	20	100.0	3212	6	AX026089	AX026089 Sequence
7	20	100.0	45121	6	AX0265635	AX0265635 Sequence
8	20	100.0	187313	9	AL157394	AL157394 Human DNA
9	19	95.0	227648	2	AC133743	AC133743 Rattus no
10	19	95.0	248752	2	AC094502	AC094502 Rattus no
11	19	95.0	264615	2	AC121457	AC121457 Rattus no
12	18.4	92.0	1454	10	RNCND1	X75207 R. norvegicu
13	18.4	92.0	1689	6	AX684916	AX684916 Sequence
14	18.4	92.0	2358	10	BC044841	BC044841 Mus muscu
15	18.4	92.0	3737	6	AX695351	AX695351 Sequence
16	18.4	92.0	3737	10	S78355	AX695351 Sequence
17	18.4	92.0	3751	10	RATCYCD1	S78355 Cyl-1-cycl
18	18.4	92.0	4300	10	MMU67925	D14014 Rattus rat
19	18.4	92.0	4329	8	SPAC38	U67925 Mus musculu
20	18.4	92.0	42861	6	AF105153	Z69086 S.pombe chr
21	18.4	92.0	48823	2	AL837517	AX695350 Sequence
22	18.4	92.0	72414	2	AC025615	AF105153 Homo sapi
23	18.4	92.0	145290	8	AC096688	AL837517 Homo sapi
24	18.4	92.0	146516	2	AC079981	AC025615 Homo sapi
25	18.4	92.0	153040	9	AC096688	AC025615 Homo sapi
26	18.4	92.0	154480	2	AC116618	AC096688 Oryza sat
27	18.4	92.0	154613	10	AC129213	AC079981 Homo sapi
28	18.4	92.0	163102	9	AC093749	AC116618 Homo sapi
29	18.4	92.0	170131	2	AC084379	AC129213 Mus muscu
30	18.4	92.0	173690	2	AC024055	AC093749 Homo sapi
31	18.4	92.0	177312	9	AC079801	AC024055 Homo sapi
32	18.4	92.0	178130	2	AC119975	AC079801 Homo sapi
33	18.4	92.0	179198	9	AC0592183	AC119975 Mus muscu
34	18.4	92.0	180141	2	AC009706	AC0592183 Human DNA
35	18.4	92.0	181997	2	AC055849	AC009706 Homo sapi
36	18.4	92.0	183417	2	AC068822	AC055849 Homo sapi
37	18.4	92.0	191190	10	AC122906	AC068822 Homo sapi
38	18.4	92.0	198073	2	AC083946	AC122906 Mus muscu
39	18.4	92.0	220190	2	AC108852	AC083946 Mus muscu
40	18.4	92.0	248187	2	AC097677	AC108852 Mus muscu
41	18.4	92.0	250008	2	AC095937	AC097677 Rattus no
42	18.4	92.0	260336	2	AC094614	AC095937 Rattus no
43	18.4	90.0	178001	10	AL772227	AC094614 Rattus no
44	18	90.0	185694	2	AC079505	AL772227 Mouse DNA
45	18	90.0	185694	2	AC079505	AC079505 Mus muscu

ALIGNMENTS

RESULT 1
AX026102
LOCUS AX026102
DEFINITION Sequence 14 from Patent DE19847779.
ACCESSION AX026102
VERSION AX026102.1 GI:10187533
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

Pred. No. is the number of results predicted by chance to have a

JOURNAL

Patent: DE 19847779-C 14 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES

Location/Qualifiers
1..20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT

4 a 2 c 3 g 11 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AATGTTGCTTAAGCTTTT 20
1 AATGTTGCTTAAGCTTTT 20

Db

RESULT 2

AX026108 40 bp DNA linear PAT 16-SEP-2000
LOCUS
DEFINITION Sequence 20 from Patent DE19847779.
ACCESSION AX026108
VERSION AX026108.1 GI:10187539
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 20 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

JOURNAL

Location/Qualifiers
1..40
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

Location/Qualifiers
1..40

BASE COUNT

7 a 8 c 6 g 19 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20
11 AATGTTGCTTAAGCTTTT 30

Db 11 AATGTTGCTTAAGCTTTT 30

RESULT 3

AX026116

LOCUS
DEFINITION Sequence 28 from Patent DE19847779.
ACCESSION AX026116
VERSION AX026116.1 GI:10187547
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 28 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

JOURNAL

Location/Qualifiers
1..40
/organism="Homo sapiens"

FEATURES

Location/Qualifiers
1..40
/organism="Homo sapiens"

BASE COUNT

7 a 8 c 6 g 19 t

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AATGTTGCTTAAGCTTTT 20
11 AATGTTGCTTAAGCTTTT 30

Db

RESULT 4

HSCD955FR 2344 bp DNA linear PRI 05-FEB-1997
LOCUS
DEFINITION H.sapiens CD95 gene 5' flanking region.
ACCESSION X87625
VERSION X87625.1 GI:902311
KEYWORDS beta interferon; CD95 gene; silencer.
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Rudert, F., Vasser, E., Forbes, L., Lindridge, E., Wang, Y. and
Watson, J.
Identification of a silencer, enhancer, and basal promoter region
in the human CD95 (Fas/Apo-1) gene
DNA Cell Biol. 14 (11), 931-937 (1995)
96069539
PUBMED 7576179

JOURNAL

Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &
Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND
Overlaps with X81335, & X82279-X82286.
Location/Qualifiers
1..2344
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/map="q24.1"
/tissue_type="Placenta"
/germline
564..1337
/note="silencer"
717..801
/note="beta interferon silencer B motifs"
complement(1237..1244)
/note="lysozyme silencer 1"
1338..1919
1920..2344
/note="basal promoter"

COMMENT

Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &
Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND
Overlaps with X81335, & X82279-X82286.
Location/Qualifiers
1..2344
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/map="q24.1"
/tissue_type="Placenta"
/germline
564..1337
/note="silencer"
717..801
/note="beta interferon silencer B motifs"
complement(1237..1244)
/note="lysozyme silencer 1"
1338..1919
1920..2344
/note="basal promoter"

FEATURES

Location/Qualifiers
1..2344

Query Match

Best Local Similarity 100.0%; Score 20; DB 9; Length 2344;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20
11 AATGTTGCTTAAGCTTTT 30

Db 11 AATGTTGCTTAAGCTTTT 30

RESULT 5

AX026092

LOCUS
DEFINITION Sequence 4 from Patent DE19847779.
ACCESSION AX026092
VERSION AX026092.1 GI:10187539
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 28 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

JOURNAL

Location/Qualifiers
1..40
/organism="Homo sapiens"

FEATURES

Location/Qualifiers
1..40
/organism="Homo sapiens"

ACCESSION AX026092 GI:10187523
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 1984779-C 4 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)
 FEATURES location/Qualifiers
 source 1..2827
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2827;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20
 437 AATGTTGCTTAAGCTTTT 456

RESULT 6
 AX026089 3212 bp DNA linear PAT 16-SEP-2000
 LOCUS
 DEFINITION Sequence 1 from Patent DE19847779.
 ACCESSION AX026089
 VERSION AX026089.1 GI:10187520
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.
 TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
 JOURNAL Patent: DE 1984779-C 1 03-FEB-2000;
 DEUTSCHES KREBSFORSCH (DE)
 FEATURES location/Qualifiers
 source 1..3212
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 778 a 784 c 809 g 841 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3212;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20
 437 AATGTTGCTTAAGCTTTT 456

RESULT 7
 AX695635 45121 bp DNA linear PAT 31-MAR-2003
 LOCUS
 DEFINITION Sequence 1262 from Patent WO03008563.
 ACCESSION AX695635
 VERSION AX695635.1 GI:29418787
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Morris, D.W. and Engelhard, E.K.
 TITLE Novel compositions and methods for cancer
 JOURNAL Patent: WO 03008583-A 1262 30-JAN-2003;
 Sagres Discovery (US)
 FEATURES location/Qualifiers
 source 1..45121
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 45121;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20
 8430 AATGTTGCTTAAGCTTTT 8449

RESULT 8
 AL157394 187313 bp DNA linear PRI 22-AUG-2001
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-399019 on chromosome 10, complete sequence.
 ACCESSION AL157394
 VERSION AL157394.15 GI:15384622
 KEYWORDS
 SOURCE HTG.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Blakey, S.
 TITLE Direct Submission
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT On Aug 31, 2001 this sequence version replaced GI:14161146. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPep; Information on the WormPep database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10> RP11-399019 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/dacpac/home.htm> VECTOR: DBAC3.6
 This sequence is the entire insert of clone RP11-399019 The true left end of clone RP11-496H23 is at 166408 in this sequence. The true right end of clone RP11-30415 is at 18704 in this sequence.
 Location/Qualifiers

FEATURES
SOURCE
1. 227648
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-182C23"
129007..130011
/note="wgs contig"
misc_feature
165940..167729
/note="wgs contig"
misc_feature
198830..199863
/note="wgs contig"
misc_feature
203077..204817
/note="wgs contig"
BASE COUNT 59974 a 38543 c 39231 g 59843 t 30057 others
ORIGIN

Query Match 95.0%; Score 19; DB 2; length 227648;
Best Local Similarity 100.0%; P-ctd No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATGCTTAACTTTT 19
15664 AATGCTTAACTTTT 15646

RESULT 10
AC094502 248752 bp DNA linear HTG 09-MAY-2003
AC094502/c
LOCUS
DEFINITION
3 unordered pieces.
AC094502 GI:30466791
AC094502.6 GI:30466791 HTGS DRAFT; HTGS_ENRICHED.
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 248752)
Wuzny,D.,Marie, Metzger,M.,Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, W., Benham, F.,
Biswal, K., Blair, D., Blankenburg, K., Blyth, P., Brown, M.,
Bryan, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chavaz, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Dalla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dim, R., Diya, K.,
Dreier, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Eaves, K.,
Egan, A., Escoto, M., Evans, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Gatta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
3 (bases 1 to 248752)
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GATW
Center clone name: CH230-4B8
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 206523 bases at least Q40
Consensus quality: 211412 bases at least Q30
Consensus quality: 214233 bases at least Q20
Consensus quality: 215006; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Kapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovari, C.,
Kovats, C., Kraft, C.B., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Luysegger, H., Lozano, R., Lu, X., Ma, J.,
Manshwar, M., Mahdantane, M., Mahmood, M., Mallory, K., Mangum, A.,
Mangum, B., Margue, P., Martin, K., Martin, R., Martinez, E.,
Mangum, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Manja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Muniesa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwankweli, O., Okunolu, G., Olatunbosun, A., Pal, S., Pankoch, C.,
Pasternak, S., Paul, H., Perez, A., Perez, A., Perz, P., Pu, L.,
Plopper, F., Poldinger, A., Popovic, D., Primus, R., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,
Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Shen, H.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Syarik, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Walker, B., Wang, J.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, A., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 245823: contig of 245823 bp in length
 * 245824 245923: gap of unknown length
 * 245924 247289: contig of 1376 bp in length
 * 247300 247399: gap of unknown length
 * 247400 248752: contig of 1353 bp in length.

FEATURES
 source
 1. 248752
 location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-488"
 1. 1697
 /note="wgs_end_extension
 clone_end:T7"
 2836. 3642
 /note="clone_boundary
 clone_end:T7
 site:ECORI
 misc_feature
 /note="wgs_end_extension
 clone_end:T7
 2836. 3642
 /note="clone_boundary
 clone_end:T7
 site:ECORI
 misc_feature
 /note="wgs_contig"
 90194. 91759
 /note="wgs_contig"
 91810. 93031
 /note="wgs_contig"
 152801. 154271
 /note="wgs_contig"
 184165. 185672
 /note="wgs_contig"
 complement(237880..238728)
 /note="clone_boundary
 clone_end:SP6
 site:ECORI
 misc_feature
 end_sequence:BH05317"
 248283. 245823
 /note="wgs_end_extension
 clone_end:SP6"
 42200 c 42265 g 64973 t 32317 others

BASE COUNT 66997 a 42200 c 42265 g 64973 t 32317 others
 ORIGIN
 Query Match 95.0%; Score 19; DB 2; Length 248752;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTCCTTAAGCTTTT 19
 Db 203668 AATGTCCTTAAGCTTTT 203650

RESULT 11
 AC121457
 LOCUS AC121457 264615 bp DNA linear HTG 12-OCT-2002
 DEFINITION Rattus norvegicus clone CH230-146J5, *** SEQUENCING IN PROGRESS
 AC121457.3 GI:23907899
 HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 1 (bases 1 to 264615)
 Muzny,D,Marie,,Metzker,M,lee,,Abramzon,S,,Adams,C,,Alder,J,,
 Allen,C,,Allen,H,,Albrooks,S,,Amin,A,,Angiano,D,,

REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Submitted (12-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 264615)
 Rat Genome Sequencing Consortium.
 Submitted (12-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 in the feature table below represents a scaffold in the Atlas
 assembly (a "contig-scaffold"). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature

Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Bismalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Butch,K., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Cente,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Gant,R., Garcia,A., Garner,T., Garza,M.,
 Gedgegoris,E., Geier,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
 Gunaratne,P., Haland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
 Hollins,B., Howells,S., Hulik,S., Hume,V., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,D., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensbawa,L., Louised,H., Lozada,R.D., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Mallory,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,S.,
 Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenan,B.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwackeleme,O., Okwono,G., Olarinbasagoun,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Frankkoc,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L.,
 Rellay,B., Reilly,M., Ren,Y., Reuter,M., Reeves,K., Reigh,M.A., Reigh,R.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Rutz,S.,
 Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Slason,I., Sitter,C.D., Smajd,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejo,Z., Uemami,K.,
 Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., Wemmi,K.,
 Williams,G., Willson,R., Wiczcyk,R., Wooden,H., Wolfe,F.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhou,D., von
 Weinstock,G., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Unpublished
 Direct Submission
 2 (bases 1 to 264615)
 Worley,K.C.
 Direct Submission
 Submitted (18-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 264615)
 Rat Genome Sequencing Consortium.
 Submitted (12-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 in the feature table below represents a scaffold in the Atlas
 assembly (a "contig-scaffold"). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYKE

Center clone name: CH230-146J5

----- Summary Statistics

Assembly program: Phrap, version 0.990329

Consensus quality: 207625 bases at least Q40

Consensus quality: 211034 bases at least Q30

Consensus quality: 213357 bases at least Q20

Estimated insert size: 215749; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 263260: contig of 263260 bp in length

* 263261: gap of unknown length

* 263361: 264615; contig of 1255 bp in length.

* Location/Qualifiers

1..264615

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-146J5"

1..1514

/note="wgs_contig"

1565..3140

/note="wgs_contig"

4867..6350

/note="wgs_contig"

complement[262294..263119]

/note="clone_boundary"

clone_end:77

site:BCCK1

end_sequence:RWBAG51TB"

BASE COUNT 63778 a 42319 c 42274 g 67234 t 49010 others

ORIGIN

Query Match 95.0%; Score 19; DB 2; Length 264615;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 19

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

FEATURES

source

gene

CDS

1..1454

/gene="CCND1"

153..1040

/gene="CCND1"

/codon_start=1

/product="CYCLIN D1"

/protein_id="CAAS3020.1"

/db_xref="GI:473123"

/db_xref="SWISS-PROT:P39948"

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

FEATURES

source

gene

CDS

1..1454

/gene="CCND1"

153..1040

/gene="CCND1"

/codon_start=1

/product="CYCLIN D1"

/protein_id="CAAS3020.1"

/db_xref="GI:473123"

/db_xref="SWISS-PROT:P39948"

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

/translation="MEHQLCCVETTPRAYPDNINLNDVRLAMLKTEETCAESVSY

FEATURES

source

gene

CDS

1..1454

/gene="CCND1"

153..1040

/gene="CCND1"

/codon_start=1</

BASE COUNT	611 a	587 c	617 g	543 t
611 a	587 c	617 g	543 t	

Search completed: August 1, 2003, 17:32:55
Job time : 84.1573 secs

Sun Aug 3 09:03:58 2003

us-09-834-291-14.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds
(without alignments)
7098.748 Million cell updates/sec

Title: US-09-834-291-14
Perfect score: 20
Sequence: 1 aatgttccttaagctttttt 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_escov:*
6: em_escpl:*
7: em_escro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	95.0	468	28	BH867204 hg93f08.y
2	19	95.0	617	29	BZ674215 PUBIT74215
3	18.4	92.0	159	9	AV243724 AV243724
4	18.4	92.0	160	9	AV174680 AV174680

5	18.4	92.0	198	9	AM213875	AM213875 uc044d03.x
6	18.4	92.0	213	9	AV134425	AV134425 AV134425
7	18.4	92.0	226	10	BB834512	BB834512 BB834512
8	18.4	92.0	237	9	AV133072	AV133072 AV133072
9	18.4	92.0	240	10	BB048825	BB048825 BB048825
10	18.4	92.0	255	10	BB546588	BB546588 BB546588
11	18.4	92.0	270	9	AV135706	AV135706 AV135706
12	18.4	92.0	278	12	B1289904	B1289904 UI-R-DKO-
13	18.4	92.0	280	10	BB049315	BB049315 BB049315
14	18.4	92.0	283	10	BF323470	BF323470 u282b07.x
15	18.4	92.0	288	9	AV219191	AV219191 AV219191
16	18.4	92.0	299	9	AV032131	AV032131 AV032131
17	18.4	92.0	328	28	AQ072724	AQ072724 HS_2255_A
18	18.4	92.0	332	9	AM229904	AM229904 uc044d03.y
19	18.4	92.0	339	10	BB546854	BB546854 BB546854
20	18.4	92.0	379	13	W78658	W78658 me88b07.r1
21	18.4	92.0	379	13	BY229322	BY229322 BY229322
22	18.4	92.0	380	13	BY174614	BY174614 BY174614
23	18.4	92.0	388	10	BB836659	BB836659 BB836659
24	18.4	92.0	411	10	BB781889	BB781889 BB781889
25	18.4	92.0	444	12	B1303131	B1303131 UI-R-DKO-
26	18.4	92.0	446	13	BY472560	BY472560 BY472560
27	18.4	92.0	457	10	BB831513	BB831513 BB831513
28	18.4	92.0	469	12	BM383441	BM383441 UI-R-DML-
29	18.4	92.0	470	12	BM898746	BM898746 UI-M-DJO-
30	18.4	92.0	491	14	CA877820	CA877820 K0960E07-
31	18.4	92.0	515	12	B1452813	B1452813 603170340
32	18.4	92.0	515	14	CA879995	CA879995 K0979F03-
33	18.4	92.0	516	14	CA872437	CA872437 K0917H06-
34	18.4	92.0	517	14	CA870300	CA870300 K0901A03-
35	18.4	92.0	518	14	CA870688	CA870688 K0904C02-
36	18.4	92.0	518	14	CA871824	CA871824 K0913E03-
37	18.4	92.0	518	14	CA873313	CA873313 K0924C08-
38	18.4	92.0	518	14	CA879757	CA879757 K0976E11-
39	18.4	92.0	527	12	BM32381	BM32381 UI-M-BH2-
40	18.4	92.0	529	9	AM259960	AM259960 um79n05.y
41	18.4	92.0	538	10	BE979816	BE979816 UI-M-BG2-
42	18.4	92.0	570	14	CA873049	CA873049 K0922E01-
43	18.4	92.0	574	12	BM386791	BM386791 UI-R-CN1-
44	18.4	92.0	577	28	AZ409434	AZ409434 1M0181X04
45	18.4	92.0	578	13	BQ510695	BQ510695 EST618110

ALIGNMENTS

RESULT 1
BH867204 468 bp DNA linear GSS 05-AUG-2002
hg93f08.y9 WGS-Zmayef (UM107 adapted methyl filtered) Zea mays
genomic clone hg93f08 5', genomic survey sequence.

DEFINITION
BH867204
GENOMIC CLONE hg93f08 5', genomic survey sequence.
ACCESSION
BH867204
VERSION
BH867204.1 GI:22103101
KEYWORDS
GSS
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 468)
Rabinowitz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,
Katzemburger, F., King, L., Miller, B., Muller, S., Nascimben, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished
Contact: W. Richard McCombie
Lila Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombie@cshl.org
Plate: hg93 row: F column: 08

TITLE
JOURNAL
COMMENT

Seq primer: -21M3UnivRev
 Class: shotgun
 High quality sequence stop: 468.
 Location/Qualifiers
 source

1. 468
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cdiv="B73"
 /db_xref="taxon:4577"
 /clone="hg93f08"
 /lab_host="JM107 or DH5a"
 /note="Organ: Immature ears; site 1: Xba I; site 2: Xba I; The vector was digested with Xba I and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was rebligated, and repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, b/g reads in pUC19). The vector ligation was transformed in either JM107 or DH5a."
 BASE COUNT 119 a 79 c 95 g 175 t
 ORIGIN

Query Match 95.0%; Score 19; DB 28; Length 468;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 2 ATGTGCTTAAGCTTTT 20
 300 ATGTGCTTAAGCTTTT 318

RESULT 2 617 bp DNA linear GSS 05-FEB-2003
 B2674215 PUBIT74TD_ZM_0.6-1.0_KB Zea mays genomic clone ZMWBTA064N04,
 LOCUS genomic survey sequence.
 ACCSSION B2674215
 VERSION B2674215.1 GI:28224015
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 1 (bases 1 to 617)
 White, C.A., Quackenbush, J., Van Aken, S., Utechtack, T., Resnick,
 A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennett, J.

AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Utechtack, T., Resnick,
 A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennett, J.
 JOURNAL Maize Genomics Consortium
 COMMENT Unpublished
 CONTACT: Cathy Whitelaw
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TP
 Class: sheared ends.

FEATURES
 source
 Location/Qualifiers
 1. 617
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMWBTA064N04"
 /clone_1lb="ZM_0.6-1.0_KB"
 /note="Vector: PCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high
 Cot selected genomic DNA library"
 BASE COUNT 166 a 103 c 235 t
 ORIGIN

Query Match 95.0%; Score 19; DB 29; Length 617;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 ATGTGCTTAAGCTTTT 20
 Db 519 ATGTGCTTAAGCTTTT 537

RESULT 3 159 bp mRNA linear EST 04-NOV-1999
 AV243724 AV243724 RIKEN full-length enriched, 0 day neonate head Mus
 musculus cDNA clone 4831423J17 3' similar to S78355 Cyl-1-cyclin D1
 mRNA sequence.
 AV243724
 AV243724.1 GI:6231183
 EST.
 Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 159)
 Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
 Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
 Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai,
 C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
 Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata,
 Suzuki, H., Takahashi, F., Tateo, M., Sugahara, Y., Suzuki, H.,
 Watanabe, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T.,
 Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Kono, H., et al. 1999)
 Unpublished

COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Sasaki, N., Izawa, M., Watanabe, M., Okazaki, Y. and Hayashizaki,
 Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki,
 Y.
 'Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Kikuchi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, D.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,
 Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for
 further details.

FEATURES
 source
 Location/Qualifiers
 1. 159
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4831423J17"
 /sex="mixed"
 /tissue-type="head"
 /dev-stage="0 day neonate"
 /lab_host="DHI0B"
 /clone_1lb="RIKEN full-length enriched, 0 day neonate
 head"
 /note="Site 1: SalI; Site 2: BamHI, cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken

DEFINITION AV134425 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA
 AV134425
 AV134425.1 GI:5320155
 EST.
 Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
 Akahira, S., Akiyama, U., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nishizawa, H., Oda, H., Owa, C.,
 Y., Suzuki, H., Suzuki, H., Tachino, M., Tomaru, Y., Tomihaga, N.,
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 RIKEN Mouse ESTs
 Unpublished
 TITLE
 JOURNAL
 COMMENT
 Contact: Chie Owa
 Genome Science Laboratory
 RIKEN
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-9145
 Fax: 81-298-36-9098
 Email: genome-res@rtc.riken.go.jp
 Thermolabile and thermostable activation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length cDNA
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3435-3460 (1998))
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for
 further details.
 FEATURES
 source
 1. 213
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="2810006111"
 /sex="mixed"
 /dev_stage="10-11 day embryo"
 /clone_id="Mus musculus C57BL/6J 10-11 day embryo"
 /clone_1id="Mus musculus C57BL/6J 10-11 day embryo"
 BASE COUNT 61 a 22 c 54 g 76 t
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 9; Length 213;
 Best Local Similarity 95.0%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AATGTCCTTAAGCTTTT 20
 Db 208 AATGTCCTTAAGCTTTT 189
 RESULT 7
 BB834512/c 226 bp mRNA linear EST 19-NOV-2001
 DEFINITION BB834512 RIKEN full-length enriched, mammary gland RCB-0527
 JYG-MC(B) cDNA Mus musculus cDNA clone G930031G12 3', mRNA
 BB834512
 BB834512.1 GI:17012755
 EST.
 Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Akimura, T., Aizawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 Hayashizaki, Y., Hiramoto, K., Hiraoka, T., Hironaka, T., Imocani, K., Ishii
 Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
 Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
 Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa
 A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, Y., Toyama
 W., Watanabe, A., Yasunishi, A., Yamamoto, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
 2001)
 TITLE
 JOURNAL
 COMMENT
 Unpublished
 Contact: Yoshinobu Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9226
 Email: genome-res@rtc.riken.go.jp
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 S., Kawai, U., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama
 Hayashizaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 FEATURES
 source
 1. 226
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G930031G12"
 /clone="G930031G12"
 /issue_type="mammary gland"
 /cell_line="RCB-0527 JYG-MC(B)"
 /clone_id="RCB-0527 JYG-MC(B) cDNA"
 /clone_1id="RCB-0527 JYG-MC(B) cDNA"
 BASE COUNT 81 a 31 c 48 g 66 t
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 10; Length 226;
 Best Local Similarity 95.0%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AATGTCCTTAAGCTTTT 20
 Db 210 AATGTCCTTAAGCTTTT 191
 RESULT 8
 AV133072/c 237 bp mRNA linear EST 01-JUL-1999
 DEFINITION AV133072 Mus musculus C57BL/6J 11-day embryo Mus musculus cDNA
 clone 2700095E14, mRNA sequence.
 AV133072
 AV133072.1 GI:5308802
 EST.
 Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,

Yano, R., Muramatsu, M. and Hayashizaki, Y.
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al.)
unpublished

indicates
, mRNA sequence.

ACCESSION BB546588
 VERSION BB546588.1 GI:9618016
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 255)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
 Izawa, M., Kadoi, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
 Kiyosawa, H., Kojima, Y., Kondo, S., Koyama, S., Kurihara, C., Kusakabe, M.,
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
 Ota, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,
 Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,
 T., Tanoda, Y., Watabiki, A., Takahashi, F., Tomioka, N., Toya, Y.,
 Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino
 RIKEN Mouse ESTs (Kono, H., et al.)

TITLE JOURNAL
 COMMENT Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki
 M., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 The mouse genome and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kikuchi, T., Akiyama, J., Shibata, K., Iwama, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, K., Iwama, M., Kawai, J.,
 Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.riken.go.jp) for
 further details.

FEATURES
 source location/Qualifiers
 1..255
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="E130310020"
 /issue_type="eyeball"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /clone_1ib="RIKEN full-length enriched, 0 day neonate
 eyeball"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer 15'
 GAGAGAGAGCGCCGCACTGAGTCTTTTCTTTTCTTTT 3', cDNA was
 transcribed using trehalose thermo-activated reverse
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence 15'
 GAGAGAGAGTCTCGAGTTATTAATATCCCTCCCTCCCT 3'. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from lambda FLC I."

BASE COUNT 89 a 31 c 57 g 78 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 255;
 Best Local Similarity 95.0%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 1 AATGTTGCTTAAGCTTTT 20
 248 AATGTTGCTTAAGCTTTT 229

RESULT 11

AV135706/c
 LOCUS AV135706 270 bp mRNA linear EST 01-JUL-1999
 DEFINITION AV135706 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA
 accession AV135706
 version AV135706.1 GI:5321436
 keywords EST.
 source Mus musculus (house mouse)
 organism Mus musculus
 comment Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Carninci, P., Shibata, K., Oza, Y., Kono, H., Itoh, M., Aizawa, K.,
 Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
 A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nishiyama, H., Oda, H., Owa, C.,
 Saito, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara
 Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomioka, N.,
 Watanabe, S., Yagawa, M., Yamamura, T., Yokota, T., Yoshino, M.,
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 RIKEN Mouse ESTs
 Unpublished
 Contact: Chie Owa
 Genome Science Laboratory
 RIKEN
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-9145
 Fax: 81-298-36-9098
 Email: genome-res@sc.riken.go.jp
 Thermofabilitation and thermoactivation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length cDNA
 (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
 Please visit our web site (http://genome.riken.go.jp) for
 further details.

FEATURES
 source location/Qualifiers
 1..270
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="2810020K08"
 /sex="mixed"
 /dev_stage="10-11 day embryo"
 /clone_1ib="Mus musculus C57BL/6J 10-11 day embryo"
 ORIGIN 88 a 34 c 61 g 87 t

Query Match 92.0%; Score 18.4; DB 9; Length 270;
 Best Local Similarity 95.0%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 1 AATGTTGCTTAAGCTTTT 20
 265 AATGTTGCTTAAGCTTTT 246

RESULT 12
 BI289904/c
 LOCUS BI289904 278 bp mRNA linear EST 19-JUL-2001

DEFINITION UI-R-DKO-cfp-f-04-0-UI.s1 UI-R-DKO Rattus norvegicus cDNA clone
 ACCESSION UI-R-DKO-cfp-f-04-0-UI 3', mRNA sequence.
 VERSION B1289904
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 278)
 AUTHORS Ronaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: Bento-soares@uiowa.edu
 Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLY=NO.

FEATURES

Source

1. 278
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-DKO-cfp-f-04-0-UI"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-DKO"
 /note="Vector: pRT13D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DKO library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRBP (20%), kidney-nRBP (20%), aorta-nRBP (20%), and placenta-nRBP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research 6: 791-806, 1996). For construction of the DKO subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (CT08), heart (CS08), kidney (CU08), aorta (CW08), and placenta (CX08). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CT0), heart (CS0), kidney (CU0), aorta (CW0), and placenta (CX0). The resulting pool of about 10,000 clones

BASE COUNT

represented about 66.6% of the final driver population.
 TAG_SEQ=None found
 107 A 41 C 58 G 72 T

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 278;
 Best Local Similarity 95.0%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 AATGTCCTTAACCTTTT 20
 |||||
 247 AATGTTCTTAACCTTTT 228

Db

RESULT 13
 BB049315/c
 LOCUS
 DEFINITION BB049315 280 bp mRNA linear EST 25-JUN-2000
 BB049315 RIKEN full-length enriched, adult male cerebellum Mus
 musculus cDNA clone 6530404H10 3' similar to S7855 Cyl-1-eyclin D1
 mRNA sequence.

ACCESSION

BB049315
 BB049315.1 GI:8456463

VERSION

BB049315
 BB049315.1 GI:8456463

KEYWORDS

EST.
 Mus musculus (house mouse)

SOURCE

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ORGANISM

1 (bases 1 to 280)
 Komu, H., Aizawa, K., Akahira, S., Akiyama, D., Arakawa, T., Carninci
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
 Hirozawa, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,
 Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya,
 T., Tsunoda, Y., Watanabe, A., Watanabe, S., Yamamura, T., Yamataka, I.,
 Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,
 M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Komu, H., et al.)

REFERENCE

Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp/
 URL: http://genome.res.riken.go.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,
 N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Thermostabilization and thermoactivation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Katsunari, T., Akiyama, U., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,
 Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.res.riken.go.jp) for
 further details.

FEATURES

location/Qualifiers
 1. 280
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="6530404H10"

```

/sex="male"
/tissue_type="cerebellum"
/dev stg="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male
cerebellum"
/notes="Site 1: XhoI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'].
GAGAGAGAGAGATCCAAAGAGCTCTCTTTTTTTTTTTTTTTT 3'}. cDNA was
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'].
GAGAGAGAGAGCGCGCCGCAATTAATTCGATTAAATTAATCCCGCCCC
3'}. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified phluescript KS(+) after bulk excision
from lambda PUC19"
BASE COUNT      90 a      40 c      59 g      91 t
ORIGIN
Query Match      92.0%; Score 18.4; DB 10; Length 280;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 AATGTTCTTAAGCTTTTTT 20
        ||||| ||||| ||||| |||||
Db      266 AATGTTCTTAAGCTTTTTT 247

RESULT 14
LOCUS      BF323470      283 bp      mRNA      linear      EST 29-DEC-2000
DEFINITION u82b07.x1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3675541 3',
mRNA sequence.
ACCESSION  BF323470
VERSION     BF323470.1
KEYWORDS   GI:11273070
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 283)
REFERENCE   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished
COMMENT    Other ESTs: u82b07.y1
            Contact: Robert Strauberg, Ph.D.
            Email: cgabds-remail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            image.lnl.gov/image/html/lresources.shtml

MGI:1436309
Seg primer: -40UP from Gibco.
FEATURES
source      1..283
            Location/Qualifiers
            1..283
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="IMAGE:3675541"
            /tissue_type="spontaneous tumor, metastatic to mammary.
            Stem cell origin."

```

```

/lab_host="DH10B"
/clone_lib="NCI_CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      87 a      57 c      40 g      99 t
ORIGIN
Query Match      92.0%; Score 18.4; DB 10; Length 283;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 AATGTTCTTAAGCTTTTTT 20
        ||||| ||||| ||||| |||||
Db      10 AATGTTCTTAAGCTTTTTT 29

RESULT 15
LOCUS      AV219191      298 bp      mRNA      linear      EST 01-NOV-1999
DEFINITION AV219191 RIKEN full-length enriched, 12 days embryo head Mus
musculus cDNA clone 3010086E05 3' similar to S78355 Cyl-1-cyclin D1
[ mice, BALB/c, brain, mRNA, 3737 nt], mRNA sequence.
ACCESSION  AV219191
VERSION     AV219191.1
KEYWORDS   GI:6168368
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 298)
REFERENCE   Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
            Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
            Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai
            C., Kawai, J., Kikuchi, N., Kojima, Y., Koyu, S., Kusakabe, M.,
            Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
            Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata
            Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H.,
            Suzuki, H., Takehashi, F., Tateo, M., Tomihaga, N., Tsunoda, Y.,
            Watanabe, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T.,
            Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
            RIKEN Mouse ESTs (Komno, H., et al. 1999)
JOURNAL    Unpublished
TITLE      Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsr.riken.go.jp,
            URL: http://genome.gsc.riken.go.jp/
            Sasaki, N., Izawa, M., Matshiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
            Matsumura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki
            Y.
            Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
            Itoh, M., Kiteunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
            Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
            Y. and Hayashizaki, Y.
            Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci, P. and Hayashizaki, Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES
source      1..298
            Location/Qualifiers
            1..298
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds
(without alignments)
8059.612 Million cell updates/sec

Title: US-09-834-291-14
Perfect score: 20
Sequence: 1 aatgtgcttaagcttttt 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AAZ88724	Human CD95 receptor
2	18.4	92.0	1689	ABN79875	Human ZBC gene se
3	16.8	84.0	514	AAF93265	Small intestine ti
4	16.8	84.0	1404	AA97038	5SKD i-antigen nuc
5	16.8	84.0	1404	AA52136	55 kDa i-antigen g
6	16.8	84.0	1410	AA97060	5SKD i-antigen cod
7	16.8	84.0	4839	AB14642	Drosophila melanog
8	16.8	84.0	140167	AB10146	Human breast cance

9	16.8	84.0	154902	24	ABQ88198	Human osteoblast d
10	16.8	84.0	165199	22	ABK83460	Human CDNA differe
11	16.4	82.0	333	22	AA560064	Human cancer agent
12	16.4	82.0	564	22	AAF93687	CDNA encoding SRT
13	16.4	82.0	1404	21	AA54576	Arabidopsis thalia
14	16.4	82.0	2613	23	AB117720	Drosophila melanog
15	16.4	82.0	3945	25	AB122326	Drosophila melanog
16	16.4	82.0	30610	22	AB115643	Human nervous syst
17	16.4	82.0	1503841	24	ABT00010	Human neuregulin 1
18	16.4	82.0	1503841	24	ABT01503	Human neuregulin-1
19	16.4	82.0	1503900	22	AAK95240	Human neuregulin-1
20	16.4	82.0	1503900	22	AAK96733	Arabidopsis thalia
21	16.4	80.0	1537	21	AA45667	Gene #3869 used to
22	16.4	80.0	1544	21	AA433610	Human immune/haema
23	16.4	80.0	6165	22	ABN97371	Human immune/haema
24	15.8	79.0	289	22	AAK76919	Human immune/haema
25	15.8	79.0	289	22	AAK76920	Human immune/haema
26	15.8	79.0	289	22	AAK76922	Human immune/haema
27	15.8	79.0	289	22	AAK76922	Human immune/haema
28	15.8	79.0	289	22	AAK81702	Human immune/haema
29	15.8	79.0	289	22	AAK81704	Human immune/haema
30	15.8	79.0	300	21	AAA00364	Human immune/haema
31	15.8	79.0	385	22	AA65104	Human colon cancer
32	15.8	79.0	413	22	AA185195	Novel human polynu
33	15.8	79.0	427	22	AA187437	Human polynucleoti
34	15.8	79.0	461	22	AAK93729	CDNA encoding SRT
35	15.8	79.0	468	24	ABN96314	Gene #2812 used to
36	15.8	79.0	468	24	ABL67114	Thyroid cancer rel
37	15.8	79.0	481	21	AA654504	Arabidopsis thalia
38	15.8	79.0	513	22	AA652226	Arabidopsis thalia
39	15.8	79.0	513	22	AA654441	Human dthp polynu
40	15.8	79.0	533	24	AAK71753	Human dthp polynu
41	15.8	79.0	628	22	AAK93449	CDNA encoding SRT
42	15.8	79.0	777	23	AA52154	Human dthp polynu
43	15.8	79.0	828	23	AA55062	Straphylococcus aur
44	15.8	79.0	828	23	AA549280	Straphylococcus aur
45	15.8	79.0	885	24	ABK73291	Bacillus lichenifo

ALIGNMENTS

RESULT 1
AAZ88724 standard; DNA: 20 BP.

AAZ88724; 11-MAY-2000 (first entry)
Human CD95 receptor promoter DNA p53 binding fragment #2.
p53; CD95 receptor; human; screening; apoptosis-modulation;
cancer chemotherapy; ss.
Homo sapiens.
DE19847779-Cl.
03-FEB-2000.
16-OCT-1998; 98DE-1047779.
16-OCT-1998; 98DE-1047779.
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
Krammer P, Mueller-Schilling M, Oren M;
WPI; 2000-162245/15.
Novel receptor DNA useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

XX Claim 2; Fig 5; 12pp; German.
XX
XX
CC This invention describes a novel p53-binding region of a human CD95
CC receptor DNA molecule. The p53-binding region, or a vector containing
CC it, can be used to screen for apoptosis-modulating substances
CC potentially useful for cancer chemotherapy. This sequence represents a
CC fragment of the human CD95 receptor promoter which is capable of
CC binding p53.
XX
SQ Sequence 20 BP; 4 A; 2 C; 3 G; 11 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATGTTGCTTAAGCTTTT 20
DB 1 AATGTTGCTTAAGCTTTT 20

RESULT 2
ABN79875/c
ID ABN79875 standard; DNA; 1689 BP.
XX
XX
AC ABN79875;
XX
XX
DT 24-JUL-2002 (first entry)
XX
XX
DE Fungal ZBC gene sequence #112.

XX
XX
KW Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein;
KW antibacterial; beta-lactam; anti-hypercholesterolaemic; lovastatin;
KW mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
KW angiogenesis inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
KW fungal toxin; cell surface receptor; plant growth regulator; pigment;
KW insecticide; antineoplastic; gene; ds.
XX
XX
OS Unidentified.

XX
XX
PN WO200224865-A2.
XX
XX
PD 28-MAR-2002.
XX
XX
PF 19-SEP-2001; 2001WO-US29288.
XX
XX
PR 19-SEP-2000; 2000US-233564P.
XX
XX
PA (MICR-) MICROBIA INC.

XX
XX
PI Holtzman D, Madden K, Maxon M, Sherman A;
XX
XX
DR WPI; 2002-352005/38.
XX
XX
DR P-PSDB; ABP35686.

XX
XX
PT New method for improving the production of a secondary metabolite e.g.
PT antineoplastic agent, ergot alkaloid from a fungus involves modulation
PT of the expression of at least one zinc binuclear cluster protein gene
XX
XX
PT Claim 2; SEQ ID 235; 49pp + sequence listing; English.

XX
XX
CC The invention relates to improving the production of a secondary
CC metabolite by a fungus. This involves modulating the expression of at
CC least one ZBC (zinc binuclear cluster protein) gene in a manner to
CC improve the yield of the secondary metabolite. Methods of the invention
CC may be used for improving the production of the secondary metabolite e.g.
CC antibacterial (such as beta-lactam), an anti-hypercholesterolaemic (such
CC as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),
CC an ergot alkaloid (such as ergotamine), an angiogenesis inhibitor (such
CC as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,
CC a fungal toxin, a modulator of cell surface receptor signalling, a plant
CC growth regulator, a pigment, an insecticide, or an antineoplastic

CC compound. The method results in a decrease in fermentor run-time, a
CC decrease in the size of the fermentor required for the production of
CC equivalent amounts of the secondary metabolite, or a decrease in the
CC biomass required for the production, which translates into decreased
CC waste that must be handled in downstream processing. The sequences given
CC in records ABN79764-ABN79911 represent ZBC genes of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 1689 BP; 508 A; 307 C; 323 G; 551 T; 0 other;
Query Match 92.0%; Score 18.4; DB 24; Length 1689;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AATGTTGCTTAAGCTTTT 20
DB 516 AATGTTGCTTAAGCTTTT 497

RESULT 3
AAF93265
ID AAF93265 standard; CDNA; 514 BP.
XX
XX
AC AAF93265;
XX
XX
DT 21-MAY-2001 (first entry)
XX
XX
DE Small intestine tissue CDNA encoding SRT protein SEQ ID 86.

XX
XX
KW Human; SRT; gene therapy; gene mapping; tissue typing; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200107611-A2.
XX
XX
PD 01-FEB-2001.
XX
XX
PF 21-JUL-2000; 2000WO-US20006.
XX
XX
PR 26-JUL-1999; 99US-0145701.

XX
XX
PA (GETH) GENENTECH INC.
XX
XX
PI Baker KP, Goddard A, Wood WI;
XX
XX
DR WPI; 2001-112729/12.

XX
XX
PT New isolated nucleic acid molecule encoding a SRT polypeptide is useful
PT for production of recombinant SRT polypeptides, gene mapping,
PT diagnosing genetic disorders and for gene therapy -
XX
XX
PT Claim 2; Fig 86; 663pp; English.

XX
XX
CC Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding
CC human SRT proteins. The CDNA sequences are isolated from various
CC different human tissue CDNA libraries. The invention relates to a method
CC for detecting CDNA encoding an SRT protein, a vector containing CDNA
CC encoding SRT, a host cell transformed with the vector, an isolated SRT
CC polypeptide, and an antibody which binds to SRT. The polynucleotide
CC sequence can be used in gene therapy and is useful in the recombinant
CC production of SRT polypeptides, as a hybridisation probe to screen
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
CC map the gene encoding the SRT polypeptides and analysing genetic
CC disorders, tissue typing and disease tissue detection. The SRT
CC polynucleotide sequences can be used in polymerase chain reaction,
CC screening for new therapeutic molecules and generation of antisense RNA
CC and DNA.

SQ Sequence 514 BP; 129 A; 111 C; 118 G; 152 T; 4 other;
Query Match 84.0%; Score 16.8; DB 22; Length 514;

Best Local Similarity 90.0%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTGCTTAAGCTTTT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 78 AATGTTGCTTAAGCTTTTCT 97

RESULT 4
AAA97038
ID AAA97038 standard; DNA; 1404 BP.

AC AAA97038;

XX 18-DEC-2000 (first entry)

DE 55KD i-antigen nucleotide sequence.

XX Immobilisation antigen; i-antigen; Ichthyophthiriasis; vaccine; ds;

KW white spot disease; freshwater fish; immune response; infection control.

XX Ichthyophthirius multifiliis.

PN WO200046373-A1.

XX 10-AUG-2000.

PF 04-FEB-2000; 2000WO-US02962.

XX 04-FEB-1999; 99US-0118634.

PR 02-MAR-1999; 99US-0122372.

PR 17-MAR-1999; 99US-0124905.

PR 27-APR-1999; 99US-0131121.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

PA (CORR.) CORNELL RES FOUND INC.

PA (CLAR.) CLARK T G.

PA (DICK.) DICKERSON H W.

PA (LINT.) LINT T.

XX Clark TG, Dickerson HW, Lint T;

PI WPI; 2000-506071/45.

XX Novel i-antigen polypeptides and polynucleotides from Ichthyophthirius

PT multifiliis, useful for prophylaxis and treatment of Ichthyophthirius

PT infection in fish

XX Claim 5; Figure 3; 14pp; English.

XX This invention relates to novel i-antigen polypeptide sequences.

XX I-antigens or immobilisation antigens are common to a variety of

XX hymenostomatid ciliates and their expression varies in response to

XX environmental stimuli. This invention relates to i-antigens in

XX Ichthyophthirius multifiliis, a protozoan which is an obligate parasite

XX of freshwater fish causing ichthyophthiriasis or white spot disease. The

XX invention includes two polypeptide and polynucleotide sequences for two

XX i-antigens, of 48 and 55 kd. Also included in the invention are

XX antibodies capable of binding to the nucleotide sequences and a method

XX for identifying i. multifiliis serotypes using the nucleotide sequences.

Query Match 84.0%; Score 16.8; DB 21; Length 1404;
Best Local Similarity 90.0%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTGCTTAAGCTTTT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 661 AATGTTGCTTAAGCTTTTCT 680

RESULT 5
AAA52136
ID AAA52136 standard; DNA; 1404 BP.

AC AAA52136;

XX 04-DEC-2000 (first entry)

DE 55 kDa i-antigen gene.

XX BTU1; beta-tubulin; protein expression system; negative selection;

KW pacitaxel sensitivity; cell surface; antigen; protozoa; ciliate;

KW live vaccine; Ichthyophthirius multifiliis; immobilization-antigen;

KW i-antigen; freshwater; fish; protozoacide; ds.

XX Ichthyophthirius multifiliis.

OS WO200046381-A1.

XX 10-AUG-2000.

PF 04-FEB-2000; 2000WO-US02966.

XX 04-FEB-1999; 99US-0118634.

PR 02-MAR-1999; 99US-0122372.

PR 17-MAR-1999; 99US-0124905.

PR 27-APR-1999; 99US-0131121.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

PA (GAER.) GAERTIG J.

PA (DICK.) DICKERSON H W.

PA (CLAR.) CLARK T G.

XX Gaertig J, Dickerson HW, Clark TG;

PI WPI; 2000-514962/46.

XX P-PSDB; AAY971177.

XX Recombinant expression systems for expressing heterologous nucleic

PT acids and producing recombinant protein, comprises nonpathogenic

PT protozoa such as Tetrahymena resistant to pacitaxel

XX Disclosure; Fig 3B; 83pp; English.

XX Tetrahymena thermophila expresses two major beta-tubulin genes (BTU1 and

XX BTU2) which encode identical beta-tubulin proteins. Either of these two

XX genes (but not both at once) can be disrupted without a detectable change

XX in the cell phenotype. A K350L substitution in the BTU1 beta-tubulin

XX protein confers increased resistance to microtubule-depolymerizing drugs

XX and increased sensitivity to pacitaxel, a microtubule-stabilizing drug.

XX Cells carrying the BTU1-K350M allele can be transformed to pacitaxel

XX resistance by gene replacement of BTU1-K350M with a wild-type BTU1 gene

XX fragment, eliminating the need to incorporate a means for positive

XX selection. Where the host organism is not a T. thermophila mutant

XX containing the BTU1-K350M allele, BTU1::neol construct, which

XX substitutes the coding region of the neol gene (conferring resistance to

CC paromycin) for that of BTUL, can be used to generate BTUL gene knockouts
 CC and for positive selection. Heterologous nucleic acids (especially
 CC encoding antigenic polypeptides) can be inserted into a BTU gene for
 CC successful cell-surface expression that is maintained by way of negative
 CC selection. Preferred expression vectors disrupt the BTUL-1K350M gene by
 CC homologous recombination-mediated insertion of a heterologous nucleic
 CC acid, thereby restoring resistance to pactitaxel in the resulting
 CC transgenic host. Transgenic ciliated protozoa are useful as live vaccines
 CC for stimulating an immune response in a vertebrate. The transgenic
 CC protozoan host cells are also useful for producing polyclonal antibodies
 CC (claimed). In particular, Tetrahymena expressing Ichthyophthirius
 CC multifiliis immunization antigen (i-antigen) protein on their surface
 CC are effective vehicles for vaccination of freshwater fish against
 CC infection by I. multifiliis.

SQ Sequence 1404 BP; 447 A; 241 C; 256 G; 460 T; 0 other;

Query Match Best Local Similarity 84.0%; Score 16.8; DB 21; Length 1404;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20
 DB 661 AATGTTGCTTAAGCTTTT 680

RESULT 6
 AAA97060

ID AAA97060 standard; DNA; 1410 BP.

XX AAA97060;

DT 18-DEC-2000 (first entry)

DE 55KD i-antigen coding region.

XX Immobilisation antigen; i-antigen; Ichthyophthiriasis; vaccine; ds;

KW white spot disease; freshwater fish; immune response; infection control.

OS Ichthyophthirius multifiliis.

XX WO200046373-A1.

PD 10-AUG-2000.

PF 04-FEB-2000; 2000WO-US02962.

XX 04-FEB-1999; 99US-0118634.

PR 02-MAR-1999; 99US-0122372.

PR 17-MAR-1999; 99US-0124905.

PR 27-APR-1999; 99US-0131121.

PA (UYGE-) UNIV GEORGIA RES FOUND INC.

PA (COOR) CORNELL RES FOUND INC.

PA (DIK/) CLARK T G.

PA (DIK/) DICKERSON H W.

PA (LINT/) LIN T.

PI Clark TG, Dickerson HW, Lin T;

DR WPI; 2000-506071/45.

PT Novel i-antigen polypeptides and polynucleotides from Ichthyophthirius
 PT multifiliis, useful for prophylaxis and treatment of Ichthyophthirius
 PT infection in fish -

PS Disclosure; Figure 2; 144pp; English.

CC This invention relates to novel i-antigen polypeptide sequences.
 CC i-antigens or immobilisation antigens are common to a variety of
 CC hymenostomatid ciliates and their expression varies in response to
 CC environmental stimuli. This invention relates to i-antigens in
 CC Ichthyophthirius multifiliis, a protozoan which is an obligate parasite

CC of freshwater fish causing Ichthyophthiriasis or white spot disease. The
 CC invention includes two polypeptide and polynucleotide sequences for two
 CC i-antigens, of 48 and 55 KD. Also included in the invention are
 CC antibodies capable of binding to the nucleotide sequences and a method
 CC for identifying I. multifiliis serotypes using the nucleotide sequences.
 CC A composition (containing the i-antigen nucleotide) capable of eliciting
 CC an immune response in fish is useful for prophylaxis, treatment or for
 CC controlling I. multifiliis infection in fish. Polynucleotide or protein
 CC vaccines comprising a portion of the amplified product encoding an
 CC antigenic i-antigen polypeptide obtained is also useful for treating or
 CC preventing I. multifiliis infection in fish. Sequences AAA97036-A97042,
 CC and AAA97060, AAA97065 and AAA97089 represent i-antigen genes and gene
 CC fragments identified in the invention. Sequences AAA97043-A97064
 CC (excluding AAA97060) and AAA97071-A97088 represent primers used in the
 CC isolation of the i-antigen gene sequences. Sequences AAB25859-B25889 and
 CC AAB25893-B25906 represent i-antigen protein and peptide sequences.

SQ Sequence 1410 BP; 449 A; 240 C; 259 G; 462 T; 0 other;

Query Match Best Local Similarity 84.0%; Score 16.8; DB 21; Length 1410;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20
 DB 661 AATGTTGCTTAAGCTTTT 680

RESULT 7
 ABL14642/c

ID ABL14642 standard; cDNA; 4839 BP.

XX ABL14642;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 38408.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-6556860/75.

DR P-PSDB; ABB70539.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Claim 1; SEQ ID NO 38408; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (AAB57757-ABB72072) and the encoded proteins

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 4839 BP; 1360 A; 1051 C; 1092 G; 1336 T; 0 other;

Query Match 84.0%; Score 16.8; DB 23; Length 4839;
Best Local Similarity 90.0%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20
2454 AATGTTGCTTAAGCTTTT 2435

RESULT 8
ABT10146/c
ID ABT10146 standard; cDNA; 140167 BP.

ABT10146;

04-DEC-2002 (first entry)

Human breast cancer associated coding sequence SEQ ID NO: 280.

Human; breast specific gene; breast cancer; differential expression;

Cyclostatic; gene therapy; gene; ss.

Homo sapiens.

WO200259271-A2.

01-AUG-2002.

25-JAN-2002; 2002WO-US02176.

25-JAN-2001; 2001US-263757P.

25-APR-2001; 2001US-286090P.

23-MAY-2001; 2001US-292517P.

(GENE-) GENE LOGIC INC.

Orr MS, Nation M, Diggans JC, Zeng W;

WPI; 2002-674803/72.

Diagnosing breast cancer in a patient comprises detecting the level of

gene expression in cell or tissue samples, where a differential gene

expression is indicative of breast cancer.

Claim 1; SEQ ID NO 280; 260bp + Sequence Listing; English.

The present invention relates to methods of diagnosing breast cancer in a

patient, which comprise detecting the level of expression in a tissue

sample of two or more genes selected from those shown in ABT09867-

ABT1112, where a differential expression of the genes indicates breast

cancer. The methods are useful in diagnosing, treating, detecting the

progression, and in monitoring treatment of breast cancer in patients.

The methods are also useful as a screening tool for agents that modulate

the onset or progression of breast cancer. The breast cancer genes may be

used as diagnostic markers for the prediction or identification of the

malignant state of breast tissue, for confirming the type and progression

of cancer, and for drug screening and assays. The present sequence is a

QY 1 AATGTTGCTTAAGCTTTT 20
DB 79855 AATGTTGCTTAAGCTTTT 79836

RESULT 9
ABQ88198
ID ABQ88198 standard; cDNA; 154902 BP.

ABQ88198;

18-SEP-2002 (first entry)

Human osteoblast differentiation related cDNA SEQ ID NO 105.

Human; osteoblast; stem cell differentiation; bone tissue deposition;

osteoporosis; osteopathic; ss.

Homo sapiens.

WO200250301-A2.

27-JUN-2002.

18-DEC-2001; 2001WO-US48276.

18-DEC-2000; 2000US-255882P.

24-APR-2001; 2001US-285691P.

(GENE-) GENE LOGIC INC.

(PROC) PROCTER & GAMBLE CO.

Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

Mertz L;

WPI; 2002-557663/59.

Use of genes and their expression profiles associated with osteoblast

differentiation for screening modulators bone formation, for diagnosing

or treating e.g. osteoporosis, or as markers for the differentiation

process.

Claim 1; SEQ ID NO 105; 78bp + Sequence Listing; English.

The invention relates to genes and their expression profiles are used

for:

(a) screening modulators of precursor stem cell differentiation into

osteoblasts, or bone tissue deposition;

(b) diagnosing abnormal deposition of bone tissue, abnormal rate of

osteoblast formation or osteoporosis; or

(c) treating or monitoring treatment of the conditions cited in (b), or

monitoring the progression of bone tissue deposition.

Specific conditions include postmenopausal osteoporosis, glucocorticoid

osteoporosis or male osteoporosis, osteopenia, osteodystrophy,

drug-induced abnormalities in bone formation or bone loss, conditions

that involve altered bone metabolism (e.g. idiopathic juvenile

osteoporosis), skeletal disease linked to breast cancer, mastocytosis,

QY 1 AATGTTGCTTAAGCTTTT 20
DB 96683 AATGTTGCTTAAGCTTTT 96702

RESULT 9
ABQ88198
ID ABQ88198 standard; cDNA; 154902 BP.

ABQ88198;

18-SEP-2002 (first entry)

Human osteoblast differentiation related cDNA SEQ ID NO 105.

Human; osteoblast; stem cell differentiation; bone tissue deposition;

osteoporosis; osteopathic; ss.

Homo sapiens.

WO200250301-A2.

27-JUN-2002.

18-DEC-2001; 2001WO-US48276.

18-DEC-2000; 2000US-255882P.

24-APR-2001; 2001US-285691P.

(GENE-) GENE LOGIC INC.

(PROC) PROCTER & GAMBLE CO.

Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

Mertz L;

WPI; 2002-557663/59.

Use of genes and their expression profiles associated with osteoblast

differentiation for screening modulators bone formation, for diagnosing

or treating e.g. osteoporosis, or as markers for the differentiation

process.

Claim 1; SEQ ID NO 105; 78bp + Sequence Listing; English.

The invention relates to genes and their expression profiles are used

for:

(a) screening modulators of precursor stem cell differentiation into

osteoblasts, or bone tissue deposition;

(b) diagnosing abnormal deposition of bone tissue, abnormal rate of

osteoblast formation or osteoporosis; or

(c) treating or monitoring treatment of the conditions cited in (b), or

monitoring the progression of bone tissue deposition.

Specific conditions include postmenopausal osteoporosis, glucocorticoid

osteoporosis or male osteoporosis, osteopenia, osteodystrophy,

drug-induced abnormalities in bone formation or bone loss, conditions

that involve altered bone metabolism (e.g. idiopathic juvenile

osteoporosis), skeletal disease linked to breast cancer, mastocytosis,

RESULT 10
 ABK83460/c
 ID ABK83460 standard; cDNA, 165199 BP.
 AC ABK83460;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #31.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW fungal infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PP 03-OCT-2001; 2001WO-US30821.
 XX
 PR 03-OCT-2000; 2000US-237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI: 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX
 PS Claim 1; SEQ ID No 31; 114pp; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,

CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 165199 BP; 48510 A; 33223 C; 34406 G; 49060 T; 0 other;
 XX
 QY Query Match 84.0%; Score 16.8; DB 24; Length 165199;
 DB Best Local Similarity 90.0%; Pred. No. 5.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AATGTCCTTAACCTTTT 20
 DB 93667 AATGTCCTTAACCAATTTT 93648
 XX
 RESULT 11
 AAS60064
 ID AAS60064 standard; cDNA, 333 BP.
 XX
 AC AAS60064;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human cancer agent-sensitive marker #65.
 XX
 KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
 KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
 KW lymphocytic leukaemia; lymphoma; plasmacytoma; reticulum cell sarcoma;
 KW Hodgkin's disease; glioma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200179556-A2.
 XX
 PD 25-OCT-2001.
 XX
 PP 13-APR-2001; 2001WO-US12132.
 XX
 PR 14-APR-2000; 2000US-197538P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lillie J, Brown JL, Bolt A, Van Huffel C;
 XX
 DR WPI: 2001-602933/68.
 XX
 PT Novel nucleic acid, used as a marker to determine the effectiveness of
 PT using TAXOL to treat cancer cell growth in individuals -
 XX
 PS Claim 1; Page 107; 527pp; English.
 XX
 CC The invention relates to 1046 novel nucleic acids which are used as
 CC markers for determining the sensitivity of a cancer cell to the
 CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
 CC they are shown to express one of the 242 sensitivity markers or the
 CC cells are shown not to express one of the 804 resistance markers.
 CC The methods can be used to determine the effectiveness of TAXOL
 CC in the treatment of cancer cell growth in an individual. The markers
 CC can be used as targets in developing anti-cancer agents such as
 CC chemotherapeutic compounds. The markers can also be used as targets in
 CC developing resistance for cancer, particularly those cancers which
 CC display resistance to agents and exhibit expression of the markers. The
 CC anticancer agents developed by the novel method can be used to treat
 CC genomic sequences corresponding to the markers, in the identification of
 CC cells or tissues which mis-express the protein. Cancers which may
 CC be targeted include carcinoma (e.g. squamous cell carcinoma),
 CC sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
 CC lymphoma, plasmacytoma, reticulum cell sarcoma, Hodgkin's disease and

CC tumours (e.g. glioma). The present sequence is one of the 1046
CC novel cancer cell markers.
XX
SQ Sequence 333 BP; 79 A; 63 C; 74 G; 93 T; 24 other;
Query Match 82.0%; Score 16.4; DB 22; Length 333;
Best Local Similarity 85.0%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 AATGTTGCTTAAAGCTTTT 20
312 AATTTTCTTGGCTTTT 331
Db
RESULT 12
AAF93687
ID AAF93687 standard; cDNA; 564 BP.
AC AAF93687;
XX
XX 21-MAY-2001 (first entry)
XX
DE cDNA encoding SRT protein isolated from MCF-7 cells SEQ ID 508.
XX
XX Human; SRT; gene therapy; gene mapping; tissue typing; ss.
XX Homo sapiens.
XX WO200107611-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000WO-US20006.
XX
XX 26-JUL-1999; 99US-0145701.
XX
XX (GETH) GENENTECH INC.
XX Baker KP, Goddard A, Wood WT;
XX WPI; 2001-112729/12.
XX
XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful
XX for production of recombinant SRT polypeptides, gene mapping,
XX diagnosing genetic disorders and for gene therapy -
XX
XX Claim 2; Fig 508; 663bp; English.
XX
XX Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding
XX human SRT proteins. The cDNA sequences are isolated from various
XX different human tissue cDNA libraries. The invention relates to a method
XX for detecting cDNA encoding an SRT protein, a vector containing cDNA
XX encoding SRT, a host cell transformed with the vector, an isolated SRT
XX polypeptide, and an antibody which binds to SRT. The polynucleotide
XX sequence can be used in gene therapy and is useful in the recombinant
XX production of SRT polypeptides, as a hybridisation probe to screen
XX libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
XX map the gene encoding the SRT polypeptides and analysing genetic
XX disorders, tissue typing and disease tissue detection. The SRT
XX polynucleotide sequences can be used in polymerase chain reaction,
XX screening for new therapeutic molecules and generation of antisense RNA
XX and DNA.
XX
XX Sequence 564 BP; 143 A; 121 C; 126 G; 170 T; 4 other;
SQ
Query Match 82.0%; Score 16.4; DB 22; Length 564;
Best Local Similarity 89.5%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
2 AATGTTGCTTAAAGCTTTT 20
128 AATTTCTGAAGCTTTT 146
Db

RESULT 13
AAC54576
ID AAC54576 standard; DNA; 1404 BP.
XX
XX AAC54576;
AC
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 78358.
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 78358.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0125548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132407.
XX 30-APR-1999; 99US-0132484.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.

```
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139818.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145090.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145193.
PR 23-JUL-1999; 99US-0145194.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145219.
PR 27-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 28-JUL-1999; 99US-0145813.
PR 28-JUL-1999; 99US-0145818.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 03-AUG-1999; 99US-0146387.
PR 03-AUG-1999; 99US-0146388.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158222.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 82.0%; Score 16.4; DB 21; Length 1404;
Best Local Similarity 94.4%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TGTGCTTAAGCTTTT 20
DB 520 TGTGCTTAAGCTTTT 537

RESULT 14
ID ABL17270 standard; DNA; 2613 BP.
XX ABL17270;
AC ABL17270;
XX
```

DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 3283.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 CC interactions -
 CC Claim 1; SEQ ID NO 3283; 21bp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC Sequence 2613 BP; 874 A; 462 C; 460 G; 817 T; 0 other;
 SQ
 Query Match 82.0%; Score 16.4; DB 23; Length 2613;
 Best Local Similarity 94.4%; Pred. No. 8.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AATGTCCTTAAGCTTTT 18
 DB 2077 AATGTCCTTAAGCTTTT 2094
 RESULT 15
 ABL22326
 ID ABL22326 standard; DNA; 3945 BP.
 AC ABL22326;
 XX 26-MAR-2002 (first entry)
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 18451.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 PF

XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 CC interactions -
 CC Claim 1; SEQ ID NO 18451; 21bp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC Sequence 3945 BP; 1117 A; 835 C; 876 G; 1117 T; 0 other;
 SQ
 Query Match 82.0%; Score 16.4; DB 23; Length 3945;
 Best Local Similarity 94.4%; Pred. No. 8.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 TGTTCCTTAAGCTTTT 20
 DB 2533 TGTTCCTTAAGCTTTT 2550
 Search completed: August 1, 2003, 13:55:21
 Job time : 10.6987 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds
(without alignments)
7955.924 Million cell updates/sec

Title: US-09-834-291-14
Perfect score: 20
Sequence: 1 aatgttgcttaagctttttc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_NEW_PUB.seq:
15: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
16: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	US-09-834-291-14	Sequence 14, Appl
2	20	100.0	40	US-09-834-291-20	Sequence 28, Appl
3	20	100.0	40	US-09-834-291-28	Sequence 4, Appl
4	20	100.0	2827	US-09-834-291-4	Sequence 1, Appl
5	20	100.0	3212	US-10-027-632-26459	Sequence 36459, A
6	17.4	87.0	474	US-10-027-632-69220	Sequence 69220, A
7	17.4	87.0	591	US-09-834-291-15	Sequence 15, Appl
8	16.8	84.0	20	US-10-027-632-72944	Sequence 72944, A
9	16.8	84.0	553	US-10-027-632-9152	Sequence 9152, Ap
10	16.8	84.0	980	US-10-027-632-250885	Sequence 250885,
11	16.8	84.0	1055	US-10-027-632-250886	Sequence 250886,
12	16.8	84.0	1055	US-09-834-975-65	Sequence 65, Appl
13	16.4	82.0	333	US-10-066-543-2878	Sequence 2878, Ap
14	16.4	82.0	549	US-10-027-632-135561	Sequence 135561,
15	16.4	82.0	602	US-10-027-632-135562	Sequence 135562,
16	16.4	82.0	602	US-10-027-632-135562	Sequence 135562,

17	16.4	82.0	785	13	US-10-027-632-157839	Sequence 157839,
18	16.4	82.0	118067	15	US-10-081-327-32	Sequence 32, Appl
19	16.4	82.0	1503841	9	US-09-795-668-1	Sequence 1, Appl
20	16.4	82.0	1503841	9	US-09-795-668-1	Sequence 1, Appl
21	16.4	82.0	1503841	10	US-09-946-807-1	Sequence 59064, A
22	16.4	82.0	1501042	13	US-10-027-632-55064	Sequence 278366,
23	16	80.0	652	13	US-10-027-632-278366	Sequence 3866, Ap
24	16	80.0	6165	10	US-09-880-107-3866	Sequence 3866, Ap
25	16	80.0	83450	11	US-09-811-469-3	Sequence 40, Appl
26	16	80.0	161852	15	US-10-081-327-40	Sequence 267598,
27	15.8	79.0	437	13	US-10-027-632-283756	Sequence 283756,
28	15.8	79.0	441	13	US-10-027-632-283756	Sequence 148, App
29	15.8	79.0	468	10	US-09-964-824A-148	Sequence 2811, Ap
30	15.8	79.0	468	10	US-09-880-107-2811	Sequence 44920, A
31	15.8	79.0	475	13	US-10-027-632-44920	Sequence 200450,
32	15.8	79.0	559	13	US-10-027-632-200451	Sequence 200451,
33	15.8	79.0	559	13	US-10-027-632-290172	Sequence 290172,
34	15.8	79.0	589	13	US-10-027-632-107268	Sequence 107268,
35	15.8	79.0	592	13	US-10-027-632-211910	Sequence 211910,
36	15.8	79.0	625	13	US-10-027-632-211911	Sequence 233945,
37	15.8	79.0	625	13	US-10-027-632-233945	Sequence 233945,
38	15.8	79.0	643	13	US-10-027-632-27458	Sequence 27458, A
39	15.8	79.0	643	13	US-10-027-632-237201	Sequence 237201,
40	15.8	79.0	663	13	US-10-027-632-237202	Sequence 237202,
41	15.8	79.0	663	13	US-10-027-632-237203	Sequence 237203,
42	15.8	79.0	667	13	US-10-027-632-290173	Sequence 290173,
43	15.8	79.0	677	13	US-10-027-632-32623	Sequence 32623, A
44	15.8	79.0	677	13	US-10-027-632-32623	Sequence 32623, A
45	15.8	79.0	677	13	US-10-027-632-32623	Sequence 32623, A

ALIGNMENTS

RESULT 1
US-09-834-291-14
Sequence 14, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834, 291
PRIOR FILING DATE: 2001-08-21
CURRENT FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-14

Query Match 100.0%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTCCTTAAGCTTTT 20
DB 1 AATGTCCTTAAGCTTTT 20

RESULT 2
US-09-834-291-20
Sequence 20, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter

```
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 40
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-20
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AATGTTGCTTAAGCTTTT 20
Db 11 AATGTTGCTTAAGCTTTT 30
```

```
RESULT 3
US-09-834-291-28
Sequence 28, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 40
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-28
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AATGTTGCTTAAGCTTTT 20
Db 11 AATGTTGCTTAAGCTTTT 30
```

```
RESULT 4
US-09-834-291-4
Sequence 4, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
```

```
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2827
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-4
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 2827;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AATGTTGCTTAAGCTTTT 20
Db 437 AATGTTGCTTAAGCTTTT 456
```

```
RESULT 5
US-09-834-291-1
Sequence 1, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: P53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3212
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-1
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 3212;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AATGTTGCTTAAGCTTTT 20
Db 437 AATGTTGCTTAAGCTTTT 456
```

```
RESULT 6
US-10-027-632-36459/c
Sequence 36459, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
```

```

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36459
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-36459
```

```
Query Match      87.0%; Score 17.4; DB 13; Length 474;
Best Local Similarity 94.7%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 ATGTTGCTTAAGCTTTT 20
Db      337 ATGTTGCTTAAGCTTTT 319
```

RESULT 7

```
US-10-027-632-69220/C
; Sequence 69220, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69220
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-69220
```

```
Query Match      87.0%; Score 17.4; DB 13; Length 591;
Best Local Similarity 94.7%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 ATGTTGCTTAAGCTTTT 20
Db      454 ATGTTGCTTAAGCTTTT 436
```

RESULT 8

```
US-09-834-291-15
; Sequence 15, Application US/09834291
; Patent No. US20020042064A1
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; APPLICANT: Muller-Schilling, Martina
; APPLICANT: Oren, Moshe
; TITLE OF INVENTION: p53 Binding Areas
; FILE REFERENCE: 4121-122
; CURRENT APPLICATION NUMBER: US/09/834,291
```

```

; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: ECT/DE99/03343
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: DE 198 47 779.1
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-834-291-15
```

```
Query Match      84.0%; Score 16.8; DB 9; Length 20;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 AATGTTGCTTAAGCTTTT 20
Db      1 AATGTTGCTTAAGCTTTT 20
```

RESULT 9

```
US-10-027-632-72944/C
; Sequence 72944, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72944
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(553)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-72944
```

```
Query Match      84.0%; Score 16.8; DB 13; Length 553;
Best Local Similarity 90.0%; Pred. No. 6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 AATGTTGCTTAAGCTTTT 20
Db      38 AATGTTGCTTAAGCTTTT 19
```

RESULT 10

```
US-10-027-632-9152/C
; Sequence 9152, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```



```

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9152
; LENGTH: 980
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-9152

```

```

Query Match      84.0%; Score 16.8; DB 13; Length 980;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

1 AATGTTGCTTAAGCTTTT 20
|||||
Db      111 AATGTTCTGAAGCTTTT 92

```

```

RESULT 11
US-10-027-632-250885/c
; Sequence 250885, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250885
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-250885

```

```

Query Match      84.0%; Score 16.8; DB 13; Length 1055;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20
|||||

```

```

Db      464 AATGTTGCTTAYGTTTTT 445

```

```

RESULT 12
US-10-027-632-250886/c
; Sequence 250886, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250886
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-250886

```

```

Query Match      84.0%; Score 16.8; DB 13; Length 1055;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

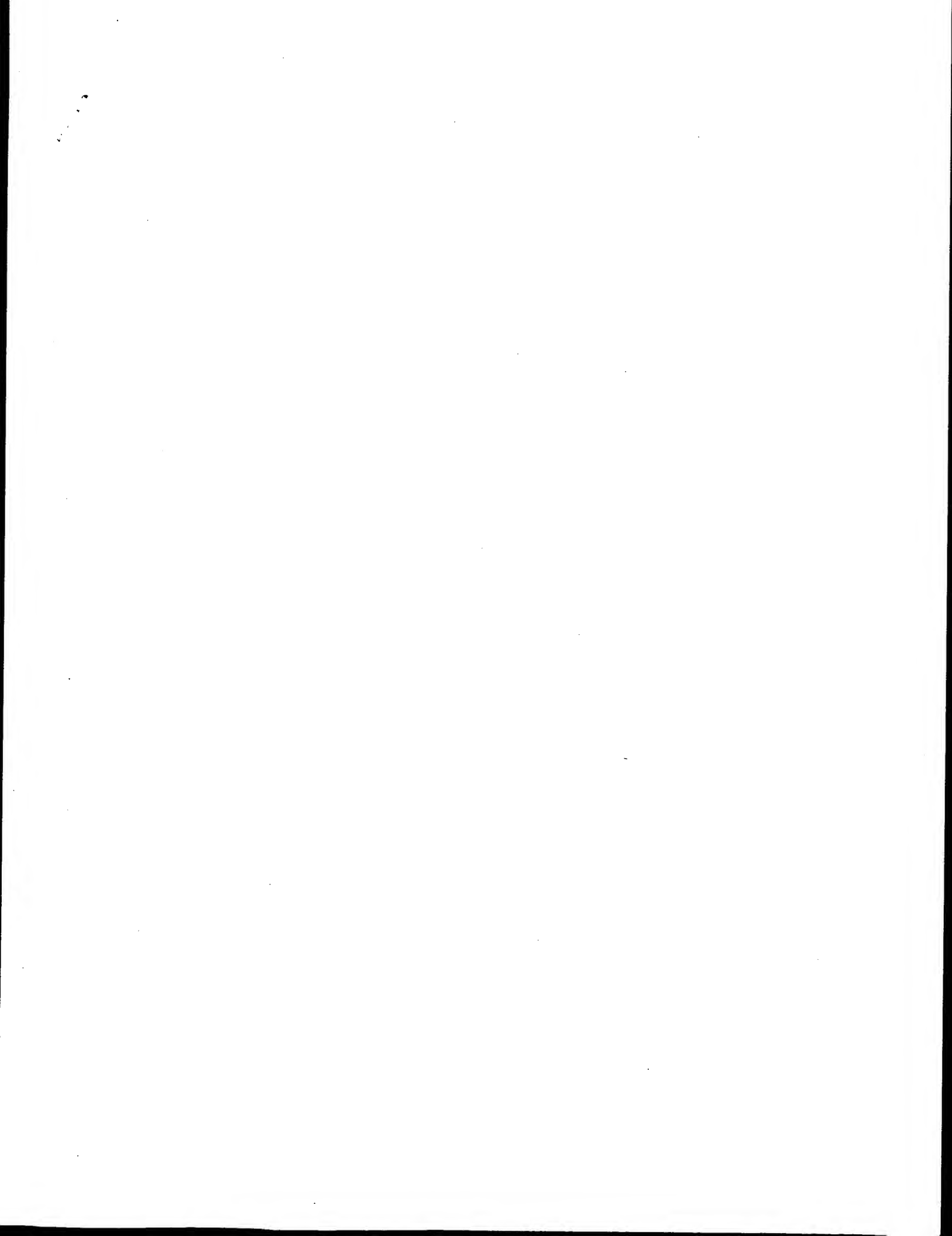
1 AATGTTGCTTAAGCTTTT 20
|||||
Db      464 AATGTTGCTTAYGTTTTT 445

```

```

RESULT 13
US-09-834-975-65
; Sequence 65, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US/09/834,975
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(333)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-65

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.959 Million cell updates/sec

Title: US-09-834-291-14

Perfect score: 20

Sequence: 1 aatgttgcttaagcttttc 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
C 1	16.8	84.0	1077 4 US-09-328-352-2249	Sequence 2249, App1
C 2	16	80.0	83450 4 US-09-811-469-3	Sequence 3, App1
C 3	15.8	79.0	186 4 US-09-328-352-3726	Sequence 3726, App
C 4	15.8	79.0	1485 4 US-09-328-352-1158	Sequence 1158, App
C 5	15.8	79.0	1551 4 US-09-328-352-2348	Sequence 2348, App
C 6	15.8	79.0	5134 2 US-08-310-912A-157	Sequence 157, App
C 7	15.8	79.0	5134 3 US-09-301-085-157	Sequence 157, App
C 8	15.8	79.0	5134 5 PCT-US95-04589-157	Sequence 157, App
C 9	15.8	79.0	5475 2 US-08-680-327-1	Sequence 1, App1
C 10	15.8	79.0	5475 3 US-08-228-246-3	Sequence 3, App1
C 11	15.8	79.0	10968 2 US-08-680-327-2	Sequence 2, App1
C 12	15.8	79.0	10968 3 US-08-228-246-1	Sequence 1, App1
C 13	15.8	79.0	580073 4 US-08-545-528D-1	Sequence 1, App1
C 14	15.8	79.0	294 4 US-08-916-421B-1	Sequence 1038, App
C 15	15.4	77.0	3024 4 US-09-016-434-1038	Sequence 534, App
C 16	15.4	76.0	521 4 US-09-620-312D-534	Sequence 214, App
C 17	15.2	76.0	521 4 US-09-643-597-214	Sequence 214, App
C 18	15.2	76.0	521 4 US-09-480-884A-214	Sequence 214, App
C 19	15.2	76.0	521 4 US-09-542-615A-214	Sequence 214, App
C 20	15.2	76.0	521 4 US-09-606-421B-214	Sequence 11, App1
C 21	15.2	76.0	602 3 US-09-040-964-11	Sequence 11, App1
C 22	15.2	76.0	602 3 US-09-123-912-11	Sequence 11, App1
C 23	15.2	76.0	602 4 US-09-643-597-11	Sequence 11, App1
C 24	15.2	76.0	602 4 US-09-480-884A-11	Sequence 11, App1
C 25	15.2	76.0	602 4 US-09-542-615A-11	Sequence 11, App1
C 26	15.2	76.0	602 4 US-09-606-421B-11	Sequence 11, App1
C 27	15.2	76.0	606 3 US-09-040-984-55	Sequence 55, App1

28	15.2	76.0	606 4 US-09-123-912-55	Sequence 55, App1
29	15.2	76.0	606 4 US-09-643-597-55	Sequence 55, App1
30	15.2	76.0	606 4 US-09-480-884A-55	Sequence 55, App1
31	15.2	76.0	606 4 US-09-542-615A-55	Sequence 55, App1
32	15.2	76.0	606 4 US-09-606-421B-55	Sequence 108, App
33	15.2	76.0	645 4 US-09-522-217-108	Sequence 1717, App
34	15.2	76.0	822 4 US-09-328-352-1177	Sequence 1, App1
35	15.2	76.0	1011 3 US-08-686-528A-1	Sequence 1, App1
36	15.2	76.0	1011 3 US-09-456-287-1	Sequence 10, App1
37	15.2	76.0	1011 4 US-08-487-429A-10	Sequence 119, App
38	15.2	76.0	1011 5 PCT-US96-05320A-119	Sequence 1995, App
39	15.2	76.0	1026 4 US-09-114-001C-1995	Sequence 101, App
40	15.2	76.0	1026 4 US-09-280-116-101	Sequence 4, App1
41	15.2	76.0	1856 4 US-08-845-258-4	Sequence 4, App1
42	15.2	76.0	1991 3 US-08-990-571-4	Sequence 4, App1
43	15.2	76.0	1991 4 US-08-723-142A-4	Sequence 4, App1
44	15.2	76.0	1991 4 US-09-528-784A-4	Sequence 4, App1
45	15.2	76.0	1991 4 US-09-559-098A-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-09-328-352-2249/c
; Sequence 2249, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2249
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2249

Query Match 84.0%; Score 16.8; DB 4; Length 1077;
Best Local Similarity 90.0%; Pred. No. 65;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 AATGTTGCTTAAGCTTTT 20
Db 616 AATGTTGCTTAAGCTTTT 597

RESULT 2
US-09-811-469-3
; Sequence 3, Application US/09811469
; Patent No. 6551809
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: CLO01171
; CURRENT APPLICATION NUMBER: US/09/811,469
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3 83450
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(83450)
; OTHER INFORMATION: n = A,T,C or G
US-09-811-469-3

Query Match 80.0%; Score 16; DB 4; Length 83450;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATGTTGCTTAAGCTTT 16
Db 62943 AATGTTGCTTAAGCTT 62958

RESULT 3
US-09-328-352-3726/c
Sequence 3726, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 3726
LENGTH: 186
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3726

Query Match 79.0%; Score 15.8; DB 4; Length 186;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AATGTTGCTTAAGCTTTT 20
Db 100 AATGTTGCTTAAGCTTTT 82

RESULT 4
US-09-328-352-1158
Sequence 1158, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 1158
LENGTH: 1485
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-1158

Query Match 79.0%; Score 15.8; DB 4; Length 1485;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AATGTTGCTTAAGCTTTT 19
Db 1085 AATGCGGCTTAAGCTTTT 1103

RESULT 5
US-09-328-352-2348/c
Sequence 2348, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 2348
LENGTH: 1551
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-2348

Query Match 79.0%; Score 15.8; DB 4; Length 1551;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AATGTTGCTTAAGCTTTT 19
Db 1492 AATGTTGCTTAAGCTTTT 1474

RESULT 6
US-08-310-912A-157/c
Sequence 157, Application US/08310912A
Patent No. 5981730
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskiewicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumitaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2904
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/310,912A
FILING DATE: September 22, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/227,360
FILING DATE: April 13, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/254001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 100254
INFORMATION FOR SEQ ID NO: 157:
SEQUENCE CHARACTERISTICS:
LENGTH: 5134 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-310-912A-157

Query Match 79.0%; Score 15.8; DB 2; Length 5134;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19
|||||
Db 462 AATGTTGATAAAGCTTTT 444

RESULT 7

US-09-301-085-157/C
Sequence 157, Application US/09301085
Patent No. 6262248
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskawicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
TITLE OF INVENTION: DETECTION METHODS
FILE REFERENCE: 00786/254002
CURRENT APPLICATION NUMBER: US/09/301,085
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: 08/310,912
EARLIER FILING DATE: 1994-09-22
EARLIER APPLICATION NUMBER: 08/227,360
EARLIER FILING DATE: 1994-04-13
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 157
LENGTH: 5134
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-301-085-157

Query Match 79.0%; Score 15.8; DB 3; Length 5134;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19
|||||
Db 462 AATGTTGATAAAGCTTTT 444

RESULT 8

PCT-US95-04589-157/C
Sequence 157, Application PC/TUS9504589
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskawicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
NUMBER OF SEQUENCES: 201
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2904
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04589

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,360
FILING DATE: 13-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/230001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 100254

INFORMATION FOR SEQ ID NO: 157:
SEQUENCE CHARACTERISTICS:
LENGTH: 5134 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-04589-157

Query Match 79.0%; Score 15.8; DB 5; Length 5134;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19
|||||
Db 462 AATGTTGATAAAGCTTTT 444

RESULT 9

US-08-680-327-1/C
Sequence 1, Application US/08680327
Patent No. 5859321
GENERAL INFORMATION:
APPLICANT: Staskawicz, Brian S., Oldroyd, Giles Edward,
APPLICANT: Salmeron, John M., Rommens, Catus
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarkist Sparkman Campbell Leigh &
ADDRESS: Winston
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
CITY: Suite 1600
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,327
FILING DATE: July 11, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/310,912
FILING DATE: September 22, 1994
CLASSIFICATION: 800
APPLICATION NUMBER: 08/227,360
FILING DATE: April 13, 1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 5151-45038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391

TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5475 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
US-08-680-327-1

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 2; Length 5475;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19
DB 2508 AATGTTGATTAAGCTTTT 2490

RESULT 10
US-09-228-246-3/c
Sequence 3, Application US/09228246
Patent No. 6245510
GENERAL INFORMATION:
APPLICANT: Staskawicz, B. S. et al.
TITLE OF INVENTION: PRF Protein and Nucleic Acid Sequences: Compositions
FILE REFERENCE: 51700
CURRENT APPLICATION NUMBER: US/09/228,246
EARLIER FILING DATE: 1999-01-11
EARLIER APPLICATION NUMBER: 08/680,327
EARLIER FILING DATE: 1996-07-11
EARLIER APPLICATION NUMBER: 08/310,912
EARLIER FILING DATE: 1994-09-22
EARLIER APPLICATION NUMBER: 08/227,360
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 5475
TYPE: DNA
ORGANISM: Lycopersicon esculentum
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (5475)
US-09-228-246-3

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 3; Length 5475;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19
DB 2508 AATGTTGATTAAGCTTTT 2490

RESULT 11
US-08-680-327-2/c
Sequence 2, Application US/08680327
Patent No. 5859321
GENERAL INFORMATION:
APPLICANT: Staskawicz, Brian S., Oldroyd, Giles Edward,
APPLICANT: Salmeron, John M., Rommens, Catus
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klatquist Sparkman Campbell Leigh &
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
CITY: Suite 1600
CITY: Portland
STATE: Oregon

COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,327
FILING DATE: July 11, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/310,912
FILING DATE: September 22, 1994
CLASSIFICATION: 800
APPLICATION NUMBER: 08/227,360
FILING DATE: April 13, 1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan, E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 5151-45038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10968 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
US-08-680-327-2

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 2; Length 10968;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19
DB 6386 AATGTTGATTAAGCTTTT 6368

RESULT 12
US-09-228-246-1/c
Sequence 1, Application US/09228246
Patent No. 6245510
GENERAL INFORMATION:
APPLICANT: Staskawicz, B. S. et al.
TITLE OF INVENTION: PRF Protein and Nucleic Acid Sequences: Compositions
FILE REFERENCE: 51700
CURRENT APPLICATION NUMBER: US/09/228,246
EARLIER FILING DATE: 1999-01-11
EARLIER APPLICATION NUMBER: 08/680,327
EARLIER FILING DATE: 1996-07-11
EARLIER APPLICATION NUMBER: 08/310,912
EARLIER FILING DATE: 1994-09-22
EARLIER APPLICATION NUMBER: 08/227,360
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 10968
TYPE: DNA
ORGANISM: Lycopersicon esculentum
FEATURE:
NAME/KEY: CDS
LOCATION: (3879) .. (8186)
FEATURE:
NAME/KEY: CDS
LOCATION: (8300) .. (9466)
US-09-228-246-1

Query Match 79.0%; Score 15.8; DB 3;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AATGTTGCTTAAGCTTTT 19
|||||
Db 6386 AATGTTGATTAAGCTTTT 6368

RESULT 13
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Query Match 79.0%; Score 15.8; DB 4;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AATGTTGCTTAAGCTTTT 20
|||||
Db 2199 AATGTTCTTAAGCTTTT 2217

RESULT 14
US-08-916-421B-1/C
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature

LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g

US-08-916-421B-1

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 4; Length 1664976;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AATGTCCTTAAGCTTTT 19
|||||
Db 789019 AATGTCCTTAAGCTTTT 789001

RESULT 15

US-09-016-434-1038

Sequence 1038, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1038:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: MUSCNOT02

CLONE: 975377

US-09-016-434-1038

Query Match 77.0%; Score 15.4; DB 4; Length 294;
Best Local Similarity 94.1%; Pred. No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AATGTCCTTAAGCTTTT 18
|||||
Db 201 AATGTCCTTAAGCTTTT 217

Search completed: August 1, 2003, 08:37:17
Job time: 5.65066 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds
(without alignments)
10468.541 Million cell updates/sec

Title: US-09-834-291-15
Perfect score: 20
Sequence: 1 aatgtcttaagattttt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl : *

1: gb_ba : *

2: gb_hcg : *

3: gb_in : *

4: gb_om : *

5: gb_ov : *

6: gb_pat : *

7: gb_ph : *

8: gb_pl : *

9: gb_pr : *

10: gb_ro : *

11: gb_stc : *

12: gb_sy : *

13: gb_un : *

14: gb_vi : *

15: em_ba : *

16: em_fun : *

17: em_hum : *

18: em_in : *

19: em_mu : *

20: em_om : *

21: em_or : *

22: em_ov : *

23: em_pat : *

24: em_ph : *

25: em_pl : *

26: em_ro : *

27: em_sts : *

28: em_un : *

29: em_vi : *

30: em_hcg_hum : *

31: em_hcg_inv : *

32: em_hcg_other : *

33: em_hcg_mus : *

34: em_hcg_pln : *

35: em_hcg_rod : *

36: em_hcg_mam : *

37: em_hcg_vrt : *

38: em_gy : *

39: em_hngo_hum : *

40: em_hngo_mus : *

41: em_hngo_other : *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6 AX026103	AX026103 Sequence
2	18.4	92.0	645	6 AR174645	AR174645 Sequence
3	18.4	92.0	1454	10 RNCND1	X75207 R.novvegicu
4	18.4	92.0	2358	10 BC044841	BC044841 Mus muscu
5	18.4	92.0	3737	6 AX655351	AX655351 Sequence
6	18.4	92.0	3737	10 S78355	S78355 Cyl-1-cycl
7	18.4	92.0	3751	10 RATCYCUD1	D14014 Rattus ratt
8	18.4	92.0	5019	3 AF060234	AF060234 Valtrimorp
9	18.4	92.0	5226	2 AC014915	AC014915 Drosophil
10	18.4	92.0	5854	3 DMHSH333	X81207 D.melanogas
11	18.4	92.0	6164	6 AX344941	AX344941 Sequence
12	18.4	92.0	16243	2 AC014426	AC014426 Sequence
13	18.4	92.0	28486	6 AX695350	AX695350 Sequence
14	18.4	92.0	30033	3 CEK10H10	Z63236 Caenorhadi
15	18.4	92.0	33930	3 CEK04H4	Z27078 Caenorhadi
16	18.4	92.0	34700	2 AC019929	AC019929 Drosophil
17	18.4	92.0	80707	2 AC016111	AC016111 Homo sapi
18	18.4	92.0	86576	2 AC112247	AC112247 Homo sapi
19	18.4	92.0	90000	2 AC068803	AC068803 Homo sapi
20	18.4	92.0	100721	5 AL591520	AL591520 Zebrafish
21	18.4	92.0	105922	2 DMHR21N6	AL121842 Drosophil
22	18.4	92.0	105922	2 AC010050	AC010050 Drosophil
23	18.4	92.0	124636	2 AC091690	AC091690 Oryza sat
24	18.4	92.0	124636	2 AP003875	AP003875 Oryza sat
25	18.4	92.0	125981	2 AC124784	AC124784 Homo sapi
26	18.4	92.0	144000	5 AL935296	AL935296 Zebrafish
27	18.4	92.0	147042	2 AC121705	AC121705 Rattus no
28	18.4	92.0	150875	2 AP005726	AP005726 Oryza sat
29	18.4	92.0	153037	2 AP005094	AP005094 Oryza sat
30	18.4	92.0	153065	2 AC092525	AC092525 Papio anu
31	18.4	92.0	153890	2 AC008380	AC008380 Homo sapi
32	18.4	92.0	153903	9 AC008503	AC008503 Homo sapi
33	18.4	92.0	153906	10 AF384675	AF384675 Mus muscu
34	18.4	92.0	154480	2 AP004565	AP004565 Oryza sat
35	18.4	92.0	155053	2 AP005523	AP005523 Oryza sat
36	18.4	92.0	155383	2 AP004396	AP004396 Oryza sat
37	18.4	92.0	158241	2 AC021804	AC021804 Homo sapi
38	18.4	92.0	162626	2 AC138551	AC138551 Danio rer
39	18.4	92.0	163280	2 AC141497	AC141497 Rattus no
40	18.4	92.0	163768	2 AC094692	AC094692 Rattus no
41	18.4	92.0	168299	3 AC113250	AC113250 Drosophil
42	18.4	92.0	168538	9 AC133565	AC133565 Homo sapi
43	18.4	92.0	168723	2 AC134719	AC134719 Rattus no
44	18.4	92.0	170280	2 AP005520	AP005520 Oryza sat

ALIGNMENTS

RESULT 1
AX026103
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX026103
Sequence 15 from Patent DE19847779.
AX026103
GI:10187534

20 bp DNA
linear PAT 16-SEP-2000

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mueller-Schilling, M., Krammer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy

JOURNAL

Patent: DE 19847779-C 15 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES

source
Location/Qualifiers
1..20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 5 a 1 c 2 g 12 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.e+03;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
1 AATGTTCTTAAGATTTT 20
Db 1 AATGTTCTTAAGATTTT 20

RESULT 2

AR174645 645 bp DNA linear PAT 17-DEC-2001
LOCUS AR174645
DEFINITION Sequence 108 from patent US 6307024.
ACCESSION AR174645
VERSION AR174645.1 GI:17914965
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 645)
Novak,J.E., Presnell,S.R., Sprecher,C.A., Foster,D.C., Holly,R.D.,
Gross,J.A., Johnson,J.V., Nelson,A.J., Dillon,S.R. and
Hammond,A.K.
Cytokine zaiaphall Ligand
Patent: US 6307024-A 108 23-OCT-2001;
Location/Qualifiers
1..645
/organism="unknown"

BASE COUNT 244 a 107 c 89 g 204 t 1 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 645;
Best Local Similarity 95.0%; Pred. No. 2.9e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
1 AATGTTCTTAAGATTTT 20
Db 383 AATGTTCTTAAGATTTT 364

RESULT 3
RNCND1/C 1454 bp mRNA linear ROD 17-FEB-1995
LOCUS RNCND1/C
DEFINITION R. norvegicus CCND1 mRNA for cyclin D1.
ACCESSION X75207
VERSION X75207.1 GI:473122
KEYWORDS CCND1 gene; cyclin D1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS 1 Bianchi,S., Fabiani,S., Muratori,M., Arnold,A., Sakaguchi,K.,
Miki,T. and Brandt,M.L.
Calcium modulates the cyclin D1 expression in a rat parathyroid
cell line
JOURNAL Biochem Biophys. Res. Commun. 204 (2), 691-700 (1994)
MEDLINE 95027382
PUBMED 7980531
REFERENCE 2 (bases 1 to 1454)
AUTHORS Brandt,M.
TITLE Direct Submission

JOURNAL

Submitted (23-SEP-1993) M. Brandt, Endocrine Unit, University of
Florence, Dept of Clinical Physiopathology, Viale Pieraccini 6,
50139 Firenze, ITALY

FEATURES

source
Location/Qualifiers
1..1454
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"

gene
CDS
1..1454
/gene="CCND1"
153..1040
/gene="CCND1"
/codon_start=1
/product="CYCLIN D1"
/protein_id="CA53020.1"
/db_xref="GI:473123"

BASE COUNT 358 a 397 c 394 g 305 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 1454;
Best Local Similarity 95.0%; Pred. No. 2.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
1 AATGTTCTTAAGATTTT 1430
Db 1449 AATGTTCTTAAGATTTT 1430

RESULT 4
BC044841/C 2358 bp mRNA linear ROD 28-JAN-2003
LOCUS BC044841/C
DEFINITION Mus musculus, cyclin D1, clone MGC:7003 IMAGE:3155470, mRNA,
complete cds.
ACCESSION BC044841
VERSION BC044841.1 GI:27924088
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 2358)
Strauberg,R.
Direct Submission

JOURNAL Submitted (23-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LMNL at: <http://image.llnl.gov>
Series: IRAK Plate: 6 Row: K Column: 9

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency OK analysis, Genomescan gene prediction, Similarity but not identity to protein.

FEATURES

source
Location/Qualifiers
1..2358
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="MGC:7003 IMAGE:3155470"
/issue_type="mammary tumor. Bcaal-/fl; MMTV-Cre model. 10 months old, gross tissue."
/clone_1ib="NCI CGAP_Mam3"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
136..1023
/codon_start=1
/product="cyclin D1"
/protein_id="AAH44841.1"
/db_xref="GI:27924089"
/translation="MEHQLCCVETIRRAYPTNLINDRVLRAMKTEKTCAPSVY
FKCQKEIVPSMKRIIVATWMLVECEKCEEVPLAMNYLDRLFLSLPKRSRLDL
GATCMFVASMKKKEIPLTAELCIYDNGIRPELLQMBELLVNKMLAAWTPDF
IEHFLSKPBEADENKQIRKRAQTFVALCATDVFTSNPMSVAAGSVAAAGNLG
SPNNFLSCYRTTHFLSRVTKCDPDLRACQEQLEALLSSLRQAQNVDPKATEEGE
VEEAGLACTPTDVRVDI"

BASE COUNT 611 a 587 c 617 g 543 t

Query Match 92.0%; Score 18.4; DB 10; Length 2358;
Best Local Similarity 95.0%; Pred. No. 2.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AATGTTCTTAAGATTTT 20
Db 1438 AATGTTCTTAAGCTTTT 1419

RESULT 5
AX695351 3737 bp DNA linear PAT 31-MAR-2003
LOCUS Sequence 978 from Patent W003008583.
DEFINITION AX695351
ACCESSION AX695351.1 GI:29418501
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Morris, D.W. and Engelhard, E.K.
Novel compositions and methods for cancer
Patent: WO 03008583-A 978 30-JAN-2003;
JOURNAL
Sagres Discovery (US)

FEATURES
source
Location/Qualifiers
1..3737
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

BASE COUNT 921 a 904 c 988 g 924 t

Query Match 92.0%; Score 18.4; DB 6; Length 3737;
Best Local Similarity 95.0%; Pred. No. 1.9e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
Db 1440 AATGTTCTTAAGCTTTT 1421

RESULT 6

S78355 3737 bp mRNA linear ROD 21-SEP-1995
LOCUS S78355
DEFINITION Cyl-1-cyclin D1 [mice, BALB/c, brain, mRNA, 3737 nt].
ACCESSION S78355
VERSION S78355.1 GI:994896

KEYWORDS
SOURCE
ORGANISM
Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 3737)
AUTHORS Smith, R., Peters, G. and Dickson, C.
TITLE Genomic organization of the mouse cyclin D1 gene (Cyl-1)
JOURNAL Genomics 25 (1), 85-92 (1995)
MEDLINE 95293413
PUBMED 7774959
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI glibdb 167715] from the original journal article. This sequence comes from Fig. 1.
Map location: 7.
Location/Qualifiers
1..3737
/organism="Mus sp."
/mol_type="mRNA"
/db_xref="taxon:10095"
1..3737
/gene="Cyl-1"
/note="cyclin D1"
138..1025
/gene="Cyl-1"
/note="this sequence comes from Fig. 1"
/codon_start=1
/product="cyclin D1"
/protein_id="AA34495.1"
/db_xref="GI:994897"

BASE COUNT 921 a 904 c 988 g 924 t

Query Match 92.0%; Score 18.4; DB 10; Length 3737;
Best Local Similarity 95.0%; Pred. No. 1.9e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AATGTTCTTAAGATTTT 20
Db 1440 AATGTTCTTAAGCTTTT 1421

RESULT 7
RATCYCLD1 3751 bp mRNA linear ROD 09-JAN-2003
LOCUS RATCYCLD1
DEFINITION Rattus rattus mRNA for cyclin D1, complete cds.
ACCESSION D14014
VERSION D14014.1 GI:413908
KEYWORDS
SOURCE
ORGANISM
Rattus rattus (black rat)
Rattus rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3751)
AUTHORS Tamura, K., Kanaoka, Y., Jimo, S., Nagata, A., Ogiso, Y., Shimizu, K.,
Hayakawa, T., Nojima, H. and Okayama, H.
TITLE Cyclin G: a new mammalian cyclin with homology to fission yeast
Cyclin G
JOURNAL Oncogene 8 (8), 2113-2118 (1993)
MEDLINE 93330551

Source
Location/Qualifiers
1. .3751

CDS

ORIGIN				
	214 C	1004 G	960 E	

1427 AATGTTTCTTAAGCTTTTTT 1408

gene, complete cds.
AF060334

AF060234.1 GI:4001823

ORGANISM *Vairimorphha necatrix*

REFERENCE
1 (bases 1 to 5019)
Burenellidae; Vairimorpha.

and Embley, T.M., Healy, B., Dorey, M.W., Doolittle, W.F.

subunit of RNA polymerase II and other proteins

MEDLINE 99110933
PUBMED 990367C
12/1 200-365 (1999)

2 (bases 1 to 5019)
Hirt, R. P. and Healy, J. C.

Direct Submission
Submitted (16-APR-1999)

Location/Qualifiers
London SW7 5BD, UK

/organism="Vairimor

```

/db_xref="taxon:6039"

```

```
/gene="RPB1"
```

/ gene="RPB1"

```

/product="largest subunit of pna polymerase"

```

GGLIDLKMGTTBAFISSCEKDNEGCDI...MEDELVIKRISIQFGLSPBEIRKSSVQIIHPETMENGFPKS

PKKLTWTKGKOILSYIIPNTIYNGKSNHEHEEDT EATTEGAAUW
 IISVNNINIEFIUSSKLINTHDDSFNGNLTTEESSNIMKILNFP
 AISY

KDEWYLOTGVCIPBIESHNT...
KXNDENLIRICISLNNNVEYNLYQKNILNLKIMGYNKIKEVISERK

452 AATCTTCTTAAGATTTTT 433

RESULT 9

LOCUS AC014915

ACCESSION AC014915 *** SEQUENCING IN PROGRESS ***

KEYWORDS HTG; HTGS PHASE2.

ORGANISM *Drosophila melanogaster*

Neoptera; Endopterygota; Diptera; Brachycera: Muscomorpha.

1 (bases 1 to 5226)

Submitted (16-NOV-1999)

COMMENT

This sequence was identified as CSW 1000000.

* NOTE: This is a 'working draft' sequence e-mail to fly@celera.com

* by the finished sequence as soon as it is available.

Source	Location/Qualifiers
1	500

```
/organism="Drosophila melanogaster"
```


Db 4643 AATGTTTCTTAAGATTTTT 4662

RESULT 12
AC014426/c 16243 bp DNA linear HTG 16-NOV-1999

LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
DEFINITION AC014426
AC014426 GI:6436909
VERSION HTG: HTGS PHASE2.
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; 1 (bases 1 to 16243)
REFERENCE Adams, M. and Venter, J.C.
AUTHORS Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT This sequence was identified as CDM:10210291 by the submitter. For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..16243
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

BASE COUNT
5707 a 3127 c 2757 g 4652 t

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 2; Length 16243;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 3760 AATGTTTCTTAAGATTTTT 3741

RESULT 13
AX695350/c 28486 bp DNA linear PAT 31-MAR-2003
LOCUS AX695350
DEFINITION Sequence 977 from Patent WO03008583.
ACCESSION AX695350
VERSION AX695350.1 GI:29418500
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE Morris, D.W. and Engelhardt, E.K.
AUTHORS Novel compositions and methods for cancer
TITLE Patent: WO 03008583-A 977 30-JAN-2003;
JOURNAL Sagres Discovery (US)
FEATURES
source
1..28486
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

BASE COUNT
6341 a 6700 c 6768 g 6414 t 2263 others

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 6; Length 28486;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 AATGTTTCTTAAGATTTTT 20

Db 15914 AATGTTTCTTAAGATTTTT 15895

RESULT 14
CEK10H10 30033 bp DNA linear INV 21-MAY-2003
LOCUS Caenorhabditis elegans cosmid K10H10, complete sequence.
DEFINITION Z83236
Z83236 GI:1729646
VERSION HTG.
KEYWORDS Caenorhabditis elegans
SOURCE Caenorhabditis elegans
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodertinae; Caenorhabditis.

REFERENCE
AUTHORS none.
TITLE Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium
JOURNAL Science 282 (5396), 2012-2018 (1998)
MEDLINE 99059613
PUBMED 9851916

REMARK The C. elegans Sequencing Consortium.
2 (bases 1 to 30033)
Percy, C.M.
AUTHORS Direct Submission
JOURNAL Submitted (09-DEC-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridgeshire CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jess@sanger.ac.uk or rwnematode.wustl.edu

COMMENT

Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.
Current sequence finishing criteria for the C. elegans genome unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is not the entire insert of clone K10H10. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
The true left end of clone K10H10 is at 1 in this sequence. The true right end of clone K10H10 is at 7962 in sequence Z83110.
The true left end of clone F57C2 is at 29930 in this sequence. The true right end of clone C13B4 is at 19748 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence Z81468.
The end of this sequence (29930..30033) overlaps with the start of sequence Z83110.
For a graphical representation of this sequence and its analysis see: <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=K10H10>
IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
location/Qualifiers
1..30033

FEATURES

source
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="II"
/clone="K10H10"
join(complement(1..212),complement(Z81468.1:23725..24037),complement(Z81468.1:22690..23349),complement(Z81468.1:21450..21734),/gene="usp-14")
join(complement(1..212),complement(Z81468.1:23725..24037),complement(Z81468.1:22690..23349),complement(Z81468.1:21450..21734))

CDS

gene

```

/gene="usp-14"
/standard_name="C13B4.2"
/notes="contains similarity to Pfam domain: PF00442
(ubiquitin carboxyl-terminal hydrolases family 2),
Score=31.3, E-value=3.1e-06, N=1, PF00443 (ubiquitin
carboxyl-terminal hydrolase family 2), Score=108.6,
E-value=4e-29, N=1
CDNA EST cm1f1 comes from this gene; CDNA EST Yk110f1.5
comes from this gene
CDNA EST Yk119a8.3 comes from this gene; CDNA EST
Yk119a8.5 comes from this gene
CDNA EST Yk143a12.3 comes from this gene; CDNA EST
Yk173f1.3 comes from this gene
CDNA EST Yk187b6.3 comes from this gene; CDNA EST
Yk143a12.5 comes from this gene
CDNA EST Yk173f1.5 comes from this gene; CDNA EST
Yk187b6.5 comes from this gene
CDNA EST Yk190b4.5 comes from this gene; CDNA EST
Yk190b4.3 comes from this gene
CDNA EST Yk455e2.5 comes from this gene; CDNA EST
Yk191g6.3 comes from this gene
CDNA EST Yk269a1.3 comes from this gene; CDNA EST
Yk282f4.3 comes from this gene
CDNA EST Yk283f6.3 comes from this gene; CDNA EST
Yk364h11.3 comes from this gene
CDNA EST Yk391b5.3 comes from this gene; CDNA EST
Yk191g6.5 comes from this gene
CDNA EST Yk269a1.5 comes from this gene; CDNA EST
Yk282f4.5 comes from this gene
CDNA EST Yk283f6.5 comes from this gene; CDNA EST
Yk364h11.5 comes from this gene
CDNA EST Yk391b5.5 comes from this gene; CDNA EST
Yk619d1.3 comes from this gene
CDNA EST Yk623a10.3 comes from this gene; CDNA EST
Yk619d1.5 comes from this gene
CDNA EST Yk623a10.5 comes from this gene; CDNA EST
Yk770e04.3 comes from this gene
CDNA EST Yk832e05.3 comes from this gene; CDNA EST
Yk832e05.5 comes from this gene
CDNA EST Yk1061a11.5 comes from this gene; CDNA EST
Yk1061a11.3 comes from this gene
CDNA EST Yk118b05.5 comes from this gene; CDNA EST
Yk118b05.3 comes from this gene
CDNA EST U3223 comes from this gene; CDNA EST Yk1278e05.3
comes from this gene
CDNA EST Yk1290c12.5 comes from this gene; CDNA EST
Yk1290c12.3 comes from this gene"
/codon_start=1
/product="C. elegans USP-14 protein (corresponding
sequence C13B4.2)"
/protein_id="CAB05785.1"
/db_xref="GI:3878545"
/db_xref="SWISS-PROT:Q17361"
/translation="MPIVAVKQEKRYVEVDTPASPMVFKALPALTQVPERKVV
IMGRLLGDDWMEGITIKENMTIMMGSGVEIPPTVLEKQANDKQAEESLALYPC
GLANGTCVENSQVMQKEVVELTKPAEEMRIREHNDRLCNALTAENSLDQRA
LRSKGRPKPRPAITLSDSPROKPEKQNEGLVMSVTRITVIGSGNITSLF
RIQETTMKCESDEVSEKRYERNQITCYVQDPRLOTGKAFPEEMRNSSELN
RDAAQKQVQISRLPKYLTVINIRPFYKESITNAKLKLSVQFPMQDLDTDCSEELK
DKLVARADIKLEBDKRLERELRKVLDEQDKLFDQDVALPTAFEDDAGSNNSEFY
VLKGIITHKGRSSQGHVAVMMRSSDDGKRLFDSEHVVVVEEA1LKTKSGGDWHSFA
VYLVEARYIKQFPELPAPVPTVEAADAPEAEVSEKQ"
join(1338..1390,1450..1576,1641..1727,1778..1997,
2359..2561,3265..3739,3923..4095)
/gene="K10H10.1"
/standard_name="K10H10.1"
/notes="contains similarity to Pfam domain: PF00083 (Sugar
and other) transporter, Score=89.7, E-value=0.00037,
N=1

```

CDS

gene

```

CDNA EST Yk503h12.5 comes from this gene
CDNA EST Yk653a11.3 comes from this gene
CDNA EST Yk653a11.5 comes from this gene"
/codon_start=1
/product="Hypothetical protein K10H10.1"
/protein_id="CAB05777.1"
/db_xref="GI:3878537"
/db_xref="GOA:Q45678"
/translation="MLKRGVEYEPPLKQIMTRASHRMWITMFSGTCLVYASRLPI
SAAAVKERRAMKTDGCVTLVSCPGVATLQVPAIRADYGAELIPYSLSMTMT
FTYHLPDRAYNTNPLVVLVLRITLQCOAPHISLAIYSKHLAADKGRFGIV
LAGSHWGTVALAGISILILWIGRRLRFQFGISILNCFVPMVDRKAGPGRSP
LPDEEVILDKKIDTIESHLAATSPCEPWPGLFRHPAFMAAIVAOYTGNSISLFP
MPSYHEHETPEPAKGVVNVPSLAVVLSIVAPVASRLSGKVTYTRKIMEGAS
LIGAIACMLMVPMTSSFMISLLIFETMAARGLHGGVSNPHDPFAMHAGSVFGVN
ACGATIGFVGYVYAGHILEATNNMSVFEVTAACVAGVAMVYTLGTGOKI1"
complement(join(4481..4583,4938..5172,5274..5949))
/gene="K10H10.2"
complement(join(4481..4583,4938..5172,5274..5949))
/standard_name="K10H10.2"
/notes="contains similarity to Pfam domain: PF00291
(Pyridoxal-phosphate dependent enzyme), Score=428.1,
E-value=2.6e-125, N=1
CDNA EST Yk115e1.5 comes from this gene; CDNA EST
Yk115e1.3 comes from this gene
CDNA EST Yk156a12.3 comes from this gene; CDNA EST
Yk173c5.3 comes from this gene
CDNA EST Yk186f5.3 comes from this gene; CDNA EST
Yk156a12.5 comes from this gene
CDNA EST Yk173c5.5 comes from this gene; CDNA EST
Yk186f5.5 comes from this gene
CDNA EST Yk290d2.3 comes from this gene; CDNA EST
Yk334g3.3 comes from this gene
CDNA EST Yk393g6.3 comes from this gene; CDNA EST
Yk418f11.3 comes from this gene
CDNA EST Yk425c10.3 comes from this gene; CDNA EST
Yk432b7.3 comes from this gene
CDNA EST Yk450f7.3 comes from this gene; CDNA EST
Yk461h2.3 comes from this gene
CDNA EST Yk466c5.3 comes from this gene; CDNA EST
Yk481e1.3 comes from this gene
CDNA EST Yk481e10.3 comes from this gene; CDNA EST
Yk410h8.3 comes from this gene
CDNA EST Yk483d3.3 comes from this gene; CDNA EST
Yk290d2.5 comes from this gene
CDNA EST Yk334g3.5 comes from this gene; CDNA EST
Yk393g6.5 comes from this gene
CDNA EST Yk418f11.5 comes from this gene; CDNA EST
Yk425c10.5 comes from this gene
CDNA EST Yk432b7.5 comes from this gene; CDNA EST
Yk461h2.5 comes from this gene
CDNA EST Yk450f7.5 comes from this gene; CDNA EST
Yk481e1.5 comes from this gene
CDNA EST Yk481e10.5 comes from this gene; CDNA EST
Yk466c5.5 comes from this gene
CDNA EST Yk410h8.5 comes from this gene; CDNA EST
Yk483d3.5 comes from this gene
CDNA EST Yk519h5.3 comes from this gene; CDNA EST

```

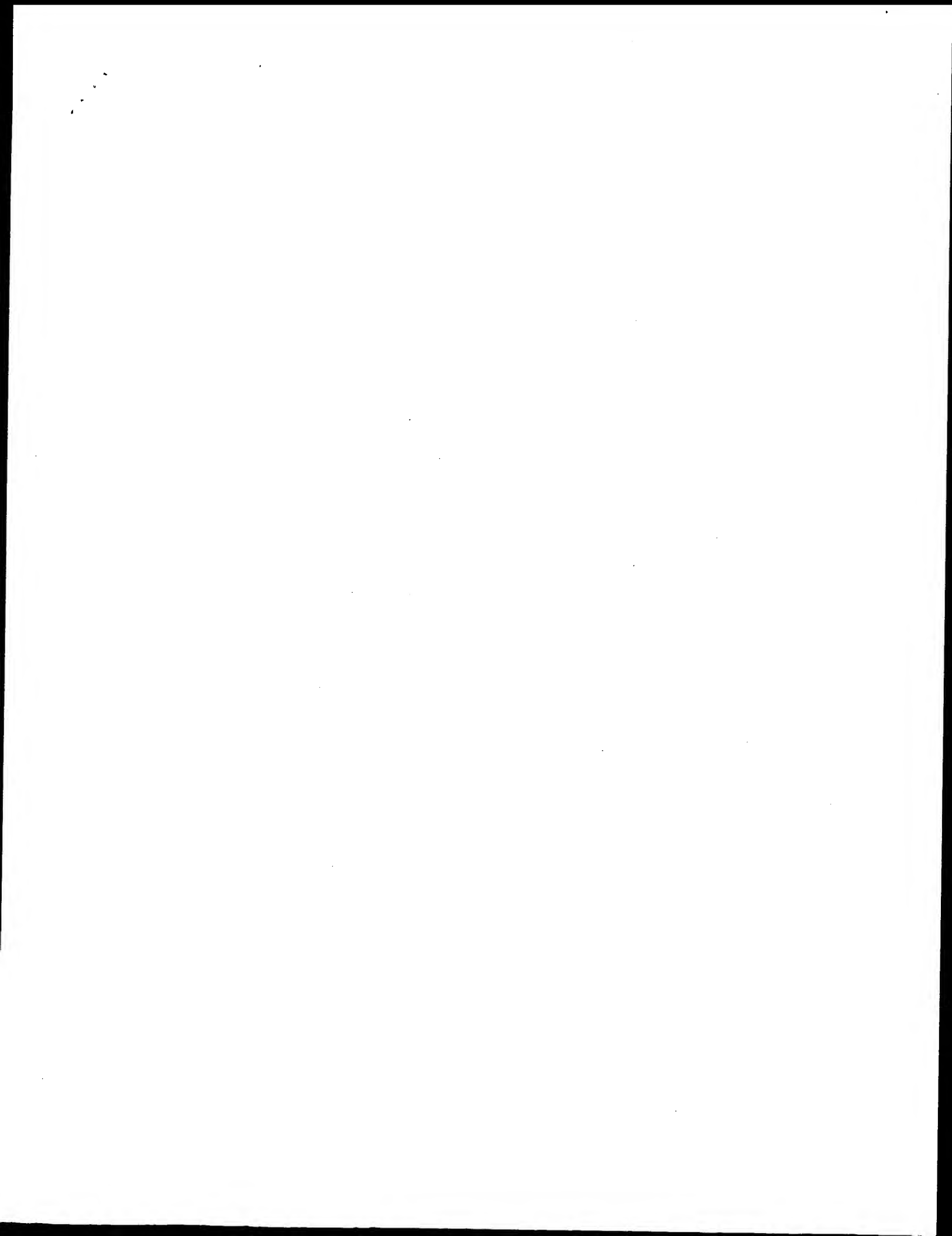
Query Match

Best Local Similarity 92.0%; Score 18.4; DB 3; Length 3003;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAGATTTTT 20
Db 14322 AATGTTCTTAATTTTT 14341

RESULT 15
CEK04H4/c 33930 bp DNA linear INV 21-MAY-2003

domain: PF01391 [collagen triple helix repeat (20
copies)] Score=669.9, E-value=4e-204, N=21; PF01413
(C-terminal tandem repeated domain in type 4 procollagen)
Score=500.5, E-value=4e-147, N=2
CDNA EST CEBSD009 comes from this gene; CDNA EST yk108f8,
comes from this gene
CDNA EST yk95b4.5 comes from this gene; CDNA EST yk95b4.3
comes from this gene
CDNA EST yk930c8.3 comes from this gene; CDNA EST yk930c8.5
comes from this gene
CDNA EST yk108f8.3 comes from this gene; CDNA EST
yk284d3.3 comes from this gene
CDNA EST yk374e1.3 comes from this gene; CDNA EST
yk425h10.3 comes from this gene
CDNA EST yk501e3.3 comes from this gene; CDNA EST
yk284d3.5 comes from this gene
CDNA EST yk374e1.5 comes from this gene; CDNA EST
yk425h10.5 comes from this gene
CDNA EST yk654a8.3 comes from this gene; CDNA EST
yk501e3.5 comes from this gene
CDNA EST yk755c11.3 comes from this gene; CDNA EST
yk755c11.5 comes from this gene
CDNA EST yk776h06.3 comes from this gene; CDNA EST
yk776h06.5 comes from this gene
CDNA EST yk878c05.3 comes from this gene; CDNA EST
yk878c05.5 comes from this gene
CDNA EST yk1029f07.3 comes from this gene; CDNA EST
yk1053c02.5 comes from this gene
CDNA EST yk1053c02.3 comes from this gene; CDNA EST
yk1062a05.5 comes from this gene
CDNA EST yk1062a05.3 comes from this gene; CDNA EST
yk1068c10.3 comes from this gene
CDNA EST yk1068f08.5 comes from this gene; CDNA EST
yk1068f08.3 comes from this gene
CDNA EST yk113c11.3 comes from this gene; CDNA EST
yk1202a01.3 comes from this gene
CDNA EST yk1278a06.5 comes from this gene; CDNA EST
yk1278a06.3 comes from this gene
CDNA EST yk1285a06.5 comes from this gene; CDNA EST
yk1285a06.3 comes from this gene
CDNA EST yk1285a05.5 comes from this gene; CDNA EST
yk1285d05.3 comes from this gene
CDNA EST yk1313a04.5 comes from this gene; CDNA EST
yk1313a04.3 comes from this gene
CDNA EST yk1025a11.3 comes from this gene"
/codon start=1
/product="C. elegans EMB-9 protein (corresponding sequence
K04H.1)"
/protein_id="CAA81584.3"
/db_xref="GI:25004987"
/db_xref="GOA:PI7139"
/translation="MSRLSLGLTAAVLISFPCDRIHYDMAACGKAPCPCTPT
KBERNGGPGGPGHGAAGDGGDGAAPGAPGAGGSGNPGKSGRGLPGSP
GHGGLGDLGLKNGERSLPCNGCDGPRMGLGAPPGQSGNNGPFGSGSPG
GHSNGQCRKLYKSGSRGRCVGLPNSGTPGLKAKADGVPGLGFGPFGSGGLKRM
GRTSGVKKEKELPGPGCGPGSGTSPMAKSTLMEVLGDPVPAQVAGVSGGLKMG
PFGMLGLGPGPYGLKQKQDGLGAGQSGSLCPNKKLFLFCRVAVERTPEIMEK
RDPKNKDLIELADTQLEGLGENDTQDQYSKETVDRKDLKEKELVDHNEILOS
TEFKAKQCDMMVAGPVGPRGVPAGAPGAPGQDGLDMGTYEGRGDRGYPGAPG
LEFGKGPEDDQCPEDBGLPGVDIGSPGLDQSGDGGPGLGDIQDREYSBKFP
GTGVKVGPPMNTGLPGRGMPKRTGVDSYPPPPNGRGRGDDCGTCDPVSNAQD
GPPNGVGPFPFGPNQDHDCCMPAPGPRKRSAGSGSPGLPAI PGVPMKAGG
ELIVGMEHPAGLPGKQHGGLGLGPRPSDGLPGVPGPGNGPGLPGVPGKPLAGD
ELKRGQSGSLGLPGLQDPRDPSPPQGTGPGYSGERGADGLPGLPGACGRGLPAPLRI
NVVQAQGLGVGMPGLPEDRSGDGLPGLPGVPDQGPPTPGMRMDGLPGFPGHGE
PGMGQDQGEVNGGLDGGCGEGLGGLGAPGAPGAPGRTGPGVQVGPQNDAG
AAGLPGPGYVNGRGLPNTGPGGAGLVNGVGLGAPGAPGRTGPGVQVGPQNDAG
PGTGQDQDQPGYSGEAGAPGQNGDGYRGLPBDQGPQSPGQDGPFGESGLPBDGLV
GPTGLRGHGNDGLPBDGLGCGESRSLGVDGPGVFGHRTGDLGLPADQDQFVGE
EAGSGRTGTVGQAPPEPGLAIPQGPQGVGPGDPGLPQDGLPGLNGRGDND
YVNPNGSLGQGPQAGLGVGDPGPGPYPGTGTGPKGSGEGLPGLPQNDQDGLRG
PPTGLGEGEGEGEGPSPQPGYPGQGGEGEGEGRTGIGENGLPGIPLQDQDQDGLRG



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds
(without alignments)
7098.748 Million cell updates/sec

Title: US-09-834-291-15
Perfect score: 20
Sequence: 1 aacgttccttaagattttt 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_hiv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_man:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: gb_gse1:*
28: gb_gse2:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	159	AV243724	AV243724 AV243724
2	18.4	92.0	160	AV174680	AV174680 AV174680
3	18.4	92.0	198	AV213875	AV213875 uc044d03.x
4	18.4	92.0	213	AV134425	AV134425 AV134425

Result No.	Score	Query Match	Length	ID	Description
5	18.4	92.0	226	BB834512	BB834512 BB834512
6	18.4	92.0	237	AV133072	AV133072 AV133072
7	18.4	92.0	240	BB048825	BB048825 BB048825
8	18.4	92.0	255	BB546588	BB546588 BB546588
9	18.4	92.0	270	AV135706	AV135706 AV135706
10	18.4	92.0	278	BI289904	BI289904 BI289904
11	18.4	92.0	280	BB049315	BB049315 BB049315
12	18.4	92.0	283	BF333470	BF333470 BF333470
13	18.4	92.0	298	AV219191	AV219191 AV219191
14	18.4	92.0	299	AV032131	AV032131 AV032131
15	18.4	92.0	328	AM027274	AM027274 AM027274
16	18.4	92.0	332	AM229904	AM229904 AM229904
17	18.4	92.0	339	BB546854	BB546854 BB546854
18	18.4	92.0	355	AM053348	AM053348 AM053348
19	18.4	92.0	360	W78658	W78658 W78658
20	18.4	92.0	379	BY229322	BY229322 BY229322
21	18.4	92.0	380	BY174614	BY174614 BY174614
22	18.4	92.0	388	BB836659	BB836659 BB836659
23	18.4	92.0	411	BB781889	BB781889 BB781889
24	18.4	92.0	436	AL719273	AL719273 AL719273
25	18.4	92.0	444	BI303131	BI303131 BI303131
26	18.4	92.0	446	BY472560	BY472560 BY472560
27	18.4	92.0	457	BB831513	BB831513 BB831513
28	18.4	92.0	470	BM758495	BM758495 BM758495
29	18.4	92.0	473	BM898746	BM898746 BM898746
30	18.4	92.0	473	BM833441	BM833441 BM833441
31	18.4	92.0	473	BM833441	BM833441 BM833441
32	18.4	92.0	491	BI452813	BI452813 BI452813
33	18.4	92.0	515	CA879995	CA879995 CA879995
34	18.4	92.0	515	CA879995	CA879995 CA879995
35	18.4	92.0	516	CA872437	CA872437 CA872437
36	18.4	92.0	517	CA870300	CA870300 CA870300
37	18.4	92.0	518	CA870300	CA870300 CA870300
38	18.4	92.0	518	CA871824	CA871824 CA871824
39	18.4	92.0	518	CA873313	CA873313 CA873313
40	18.4	92.0	518	CA879757	CA879757 CA879757
41	18.4	92.0	529	AM259960	AM259960 AM259960
42	18.4	92.0	536	AO730889	AO730889 AO730889
43	18.4	92.0	570	CA873049	CA873049 CA873049
44	18.4	92.0	572	BZ175552	BZ175552 BZ175552
45	18.4	92.0	572	BZ611328	BZ611328 BZ611328

ALIGNMENTS

RESULT 1
AV243724/c
LOCUS
DEFINITION AV243724 RIKEN full-length enriched, 0 day neonate head Mus musculus cDNA clone 4831423J17 3' similar to S78355 Cyl-1-cyclin D1 mRNA sequence.

ACCESSION AV243724
VERSION AV243724.1 GI:6231183

KEYWORDS
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.

1 (bases 1 to 159)
Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Carinci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suganara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tatenno, M., Tomimaga, N., Tsunoda, T., Watanabe, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saitoh-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Sasaki, N., Izawa, M., Matsuki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7): 3455-3460 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5): 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source location/Qualifiers

1.159
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4831423J17"
/sex="mixed"
/tissue_type="head"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_1ib="RIKEN full-length enriched, 0 day neonate head"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTCTTTT 3'], cDNA was transcribed by using trehalose thermo-activated reverse cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATTAATATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PUC I."

BASE COUNT
ORIGIN 50 a 20 c 29 g 60 t

Query Match 92.0%; Score 18.4; DB 9; Length 159;
Best Local Similarity 95.0%; Pred. No. 1.6e+04;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AATGTTCTTAAGATTTT 20
Db 152 AATGTTCTTAAGCTTTT 133

RESULT 2
LOCUS AV174680/c 160 bp mRNA linear EST 06-JUL-1999
DEFINITION AV174680 Mus musculus C57BL/6J 8-day embryo Mus musculus cDNA clone
ACCESSION AV174680
VERSION AV174680
KEYWORDS EST.
GI:5381118

SOURCE
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 160)

REFERENCE
AUTHORS Carninci, P., Shibata, K., Ozawa, Y., Kono, H., Itoh, M., Aizawa, K., Akiyama, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Horii, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Toming, N., Muramatsu, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished

TITLE
JOURNAL
COMMENT Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp

Thermosensitization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source location/Qualifiers

1.160
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="573041118"
/sex="mixed"
/dev_stage="8-day embryo"
/clone_1ib="Mus musculus C57BL/6J 8-day embryo"
/clone_2ib="Mus musculus C57BL/6J 8-day embryo"
/clone_3ib="Mus musculus C57BL/6J 8-day embryo"

BASE COUNT
ORIGIN 40 a 27 c 34 g 57 t

Query Match 92.0%; Score 18.4; DB 9; Length 160;
Best Local Similarity 95.0%; Pred. No. 1.6e+04;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AATGTTCTTAAGATTTT 20
Db 155 AATGTTCTTAAGCTTTT 136

RESULT 3
LOCUS AW213875 198 bp mRNA linear EST 03-DEC-1999
DEFINITION uc44d03.x1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:2645381 3', mRNA sequence.

ACCESSION AW213875
VERSION AW213875.1 GI:6520157
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 198)
TITLE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT Unpublished
Other ESTs: uc44d03.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNI at:
www-bio.illn.gov/bbrp/image/image.html

MG1:1025833
 Seq primer: -40UP from Gibco
 High quality sequence stop: 191.
 Location/Qualifiers

1.198

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:2645381"
 /issue_type="spontaneous tumor, metastatic to mammary."
 Stem cell origin."
 /lab_host="DH10B"
 /clone_lib="NCI-CGAP Lu29"
 /note="Organ: lung; Vector: pCMV-Sport6; Site: 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 56 a 38 c 23 g 81 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 198;
 Best Local Similarity 95.0%; Pred. No. 1.5e+04;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
 |||||
 33 AATGTTCTTAAGCTTTT 52

Db

RESULT 4
 AV134425/c 213 bp mRNA linear EST 01-JUL-1999
 LOCUS AV134425 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA
 DEFINITION clone 2810006111, mRNA sequence.
 AV134425
 AV134425.1 GI:5320155

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 213)
 Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
 Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
 A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nishizuma, H., Oda, H., Owa, C.,
 Sato, K., Shibata, Y., Shigenoto, Y., Shiraki, T., Sogabe, Y., Sugahara
 Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Tomioka, N.,
 Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE
 JOURNAL
 COMMENT

Unpublished
 Contact: Chie Owa
 Genome Science Laboratory
 RIKEN
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-9145
 Fax: 81-298-36-9098
 Email: genome-res@cc.riken.go.jp
 Thermolabile enzymes by
 trehalose and its application for the synthesis of full length cDNA
 (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for
 further details.

FEATURES
 source

Location/Qualifiers

1.213
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="2810006111"
 /sex="mixed"
 /dev_stage="10-11 day embryo"
 /clone_lib="Mus musculus C57BL/6J 10-11 day embryo"

BASE COUNT 61 a 22 c 54 g 76 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 213;
 Best Local Similarity 95.0%; Pred. No. 1.5e+04;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
 |||||
 208 AATGTTCTTAAGCTTTT 199

Db

RESULT 5
 BB834512/c 226 bp mRNA linear EST 19-NOV-2001
 LOCUS BB834512 RIKEN full-length enriched, mammary gland RCB-0527
 DEFINITION JY9-MC(B) cDNA Mus musculus cDNA clone G930031G12 3', mRNA
 sequence.
 BB834512
 BB834512.1 GI:17012755

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 226)
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imocani, K., Ishi
 Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,
 Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R., Okazaki, Y., Okido, T.,
 Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
 Shibata, K., Shigenoto, Y., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa
 A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyo, T.,
 Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE
 JOURNAL
 COMMENT

Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@cc.riken.go.jp,
 url: <http://genome-gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanabe, S., Yagame, M., Ishikawa, T., Owa, K., Tanaka, T., Matsuyama
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..226

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="G930031G12"

/tissue_type="mammary gland"

/cell_line="RCB-0527 Jyg-MC(B)"

/clone_lib="RCB-0527 Jyg-MC(B) cDNA"

RCB-0527 Jyg-MC(B) cDNA

81 a 31 c 48 g 66 c

Query Match 92.0%; Score 18.4; DB 10; Length 226;

Best Local Similarity 95.0%; Pred. NO. 1.5e+04;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AATGTTCTTAAGATTTT 20

210 AATGTTCTTAAGATTTT 191

237 bp mRNA linear EST 01-JUN-1999

AV133072 Mus musculus C57BL/6J 11-day embryo Mus musculus CDNA

AV133072

AV133072.1 GI:5308802

EST

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 237)

Carninci, P., Shibata, K., Ozawa, Y., Komno, H., Itoh, M., Aizawa, K.,

Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara

Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuna, H., Oda, H., Owa, C.,

Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara

Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N.,

Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshida, K., Yoshida, K.,

Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

RIKEN Mouse ESTs

Unpublished

Contact: Chie Owa

Genome Science Laboratory

RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-9145

Fax: 81-298-36-9098

Email: genome-res@gs.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length CDNA

(Proc. Natl. Acad. Sci. U.S.A. 95(12):520-524 (1998))

Transcriptional sequencing: A method for DNA sequencing using RNA

polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))

Please visit our web site (<http://genome.rtc.riken.go.jp>) for

further details.

Location/Qualifiers

1..237

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="2700095E14"

/sex="mixed"

/dev_stage="11-day embryo"

/clone_lib="Mus musculus C57BL/6J 11-day embryo"

83 a 27 c 49 g 78 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 237;

Best Local Similarity 95.0%; Pred. NO. 1.5e+04;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AATGTTCTTAAGATTTT 20

232 AATGTTCTTAAGATTTT 213

240 bp mRNA linear EST 25-JUN-2000

BB048825 RIKEN full-length enriched, adult male olfactory bulb Mus

musculus CDNA clone 643056J05 3' similar to S7835 Cyl-1-cyclin D1

, mRNA sequence.

BB048825

BB048825.1 GI:8455973

EST

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 240)

Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci

P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,

Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,

Kiyosawa, H., Kojima, Y., Kondo, S., Koyu, S., Kurikawa, C., Kusakabe, M.,

Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,

Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,

Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,

Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya

T., Tsunoda, Y., Watanabe, S., Watanabe, S., Yamamura, T., Yamataka, I.,

Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshida, K., Yoshida, K.,

Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshida, K., Yoshida, K.,

Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Komno, H., et al.)

Unpublished

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki

N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermostabilization and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95(12):520-524 (1998)

Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki

Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5):463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for

further details.

Location/Qualifiers

1..240

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="643056J05"

/sex="male"

/tissue_type="olfactory brain"

/dev_stage="adult"

COMMENT

Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rc.riken.go.jp
Thermostabilization and thermostabilization of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers
1..270
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2810020K08"
/sex="mixed"
/dev_stage="10-11 day embryo"
/clone_id="Mus musculus C57BL/6J 10-11 day embryo"
34 c 61 g 87 t

BASE COUNT

88 a 34 c 61 g 87 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 270;
Best Local Similarity 95.0%; Pred. No. 1.5e+04;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 AATGTTCTTAAGATTTT 20
|||||
265 AATGTTCTTAAGCTTTT 246

Db

RESULT 10

BI289904/c 278 bp mRNA linear EST 19-JUN-2001
UI-R-DKO-cfp-f-04-0-UI.s1 UI-R-DKO Rattus norvegicus cDNA clone

LOCUS

DEFINITION BI289904 UI-R-DKO-cfp-f-04-0-UI.s1 UI-R-DKO Rattus norvegicus cDNA clone
BI289904
VERSION BI289904.1 GI:14947954
KEYWORDS EST.

ACCESSION

KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus

SOURCE

Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 278)
Bonald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL

GENOME RESEARCH
MEDLINE 97044477
PUBMED 8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

COMMENT

Email: bento-soares@uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Research Genetics (www.resgen.com)
Seq Primer: M13 Forward
POLY-A-NO.

FEATURES

source

Location/Qualifiers
1..278
/organism="Rattus norvegicus"
/mol_type="mRNA"

/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DKO-cfp-f-04-0-UI"
/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"
/clone_id="UI-R-DKO"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-DKO
library is a subtracted library derived from a mixture of
five individually tagged normalized rat libraries:
brain-nRP (20%), heart-nRP (20%), kidney-nRP (20%),
aorta-nRP (20%), and placenta-nRP (20%). Each original
library was constructed from a mixture of equal amounts of
RNA from seven different developmental time-points:
embryonic day 17, embryonic day 19, embryonic day 21,
adult day 1, adult day 12, adult day 75, and adult day
200. (Exception: the aorta pool does not contain embryonic
day 17 RNA and the placenta pool contains only the three
embryonic stages). Each library was normalized
individually according to the procedure described by
Bonald, Lennon & Soares (Genome Research 6:
791-806, 1996). For construction of the DKO subtracted
library, plasmid DNA from each of the five individually
tagged normalized libraries was mixed in the proportions
specified above and electroporated into competent bacteria
for production of single-stranded circular DNA
representing the pool of libraries. Single-stranded
circular DNA representing these five normalized libraries
was then used as a tracer in a subtractive hybridization
template preparation) comprising: a) a set of about 1,000
arrayed clones from each of the five non-normalized
libraries of brain (CTOs), heart (CSOs), kidney (CUOs),
aorta (CWOs), and placenta (CXOs). The resulting pool of
approximately 5,000 clones represented about 33.3% of the
final driver population. A set of about 2,000 arrayed
clones from each of the five normalized libraries of brain
(CTO), heart (CSO), kidney (CUO), aorta (CWO), and
placenta (CXO). The resulting pool of about 10,000 clones
represented about 66.6% of the final driver population.
TAG_SEQ=None found"

BASE COUNT 107 a 41 c 58 g 72 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 278;
Best Local Similarity 95.0%; Pred. No. 1.5e+04;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 AATGTTCTTAAGATTTT 20
|||||
247 AATGTTCTTAAGCTTTT 228

Db

RESULT 11

BB049315/c 280 bp mRNA linear EST 25-JUN-2000
BB049315 RIKEN full-length enriched, adult male cerebellum Mus
musculus cDNA clone 6530404H10 3' similar to 57835 Cyl-1-cyclin D1
mRNA sequence.

LOCUS

DEFINITION BB049315 RIKEN full-length enriched, adult male cerebellum Mus
musculus cDNA clone 6530404H10 3' similar to 57835 Cyl-1-cyclin D1
mRNA sequence.

ACCESSION

VERSION BB049315.1 GI:8456463
KEYWORDS EST.

SOURCE

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
1 (bases 1 to 280)

REFERENCE

1
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadoya, K., Kagawa, I., Kai, J., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Maki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,

REFERENCE	1 (bases 1 to 298)
AUTHORS	Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Hirazane, T., Hori, F., Iwatsuki, Y., Hara, A., Hayatsu, N.

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-reesgsc.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>
Sasaki, N., Izawa, M., Matabiki, M., Ozuwa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95 (7), 3455-3460 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, T., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

Source

```
location/Qualifiers
1..298
/organism="Mus musculus"
/mol_type="cRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="3010086E05"
/sex="mixed"
/tissue_type="head"
/dev_stage="12 days embryo"
/lab_host="SOLR"
/clone_id="RIKEN full-length enriched, 12 days embryo"
head
```

BASE COUNT
ORIGIN

	Query Match	Score 18.4;	DB 9;	Length 298;
	Best Local Similarity	95.0%;	Pred. No. 1.5e+04;	
	Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;
QY	1 AATGTTCTTACGATTTT	20		
DB	293 AATTTCTTACGCTTTT	274		

RESULT 14	AV032131/c	299 bp	mRNA	linear	EST 31-AUG-1999
LOCUS	AV032131				
DEFINITION	AV032131 Mus musculus adult C57BL/6J cerebellum Mus				
ACCESSION	AV032131				
VERSION	AV032131				
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				

REFERENCE
AUTHORS
(pages 1 to 299)
Carninci, P., Shibata, K., Ozawa, Y., Komu, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, U., Kikuchi, N., Kojima, Y., Matsuyama, T., Mitsuuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shitaki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomioka, N., Muranabe, S., Yagame, M., Yamamoto, T., Yokota, T., Yoshino, M., Watanabe, K., Okazaki, Y. and Hayashizaki, Y.
RICKER NEWS ESTS
Unpublished
Contact: Chie Owa
JOURNAL
COMMENT

Source

RIKEN
Genetic Information Laboratory
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel.: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
thermosensitization and thermoactivation of thermolabile enzymes by
trehalose and its application for the syntheses of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

BASE COUNT
ORIGIN

	92.0%;	Score 18.4;	DB 9;	Length 299;
Query March	Best Local Similarity	95.0%;	Pred. No. 1.5e+04;	
Matches 19;	Conservative	0;	Mismatches 1;	Indels 0; Gaps .0.
QY	1 AATGTTCTTAGAGATTTC	20		
Dd	AATGTTCTTAGAGTTTC	275		

RESULT 15

LOCUS	A0072724	328 bp	DNA
DEFINITION	HS_2255 AI D10 MR CIT Approved Human Genomic Sperm library D Homo sapiens genomic clone Plate=2255 Col=19 Row=G, genomic survey sequence.		
ACCESSION	A0072724		
VERSION	A0072724.1	GI:3388927	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE AUTHORS

REFERENCE AUTHORS

1 (bases 1 to 328)
Mahairas, G. G., Wallace, J. C., Smith, K., Swartzell, S., Holzmann, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M. D., and
Hood, L.

TITLE

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL
MEDLINE

PROC. NATL. ACAD. SCI. U.S.A. 96 (17), 9739-9744 (1999)
99380589

**MEDLINE
PUBMED**

99380589
10449764

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington

University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA

File: (206) 616-3618
Fax: (206) 616-3887

Fax: (202) 616-3867
Email: jiwallace@washington.edu

Sequence Tagged Contigs:

Plate: 2255 Row: C

Class: BAC ends

High quality sequence stop: 328

Location/Qualifiers

1. .328

FEATURES

```
Location/Qualifiers
1. .328
/organism="Homo sapiens"
/mol_type="genomic DNA"
```

```
/db xref="taxon:9606"
```

```
/clone="Plate=2255 Col=
```

```
/sex="male"
```

```
/clone_lib="CIT Approve
```

```
/note="Organ: sperm; Ve
```

E-COL1 DH10B

BASE COUNT	102 a	63 c	50 g	1
ORIGIN				

ORIGIN

Query Match 92.0%; Score 18.4

Best Local Similarity	95.0%;	Pred. No.
-----------------------	--------	-----------

Matches 19; Conservative 0; Mismatch

.....

QY 1 AATGTTCTTAAGATTTT 20

[illegible]

06 1111111111111111 / 2

00

Search completed: August 1, 2003, 20:43:07
Job time : 69.4754 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds
(without alignments)
8059.612 Million cell updates/sec

Title: US-09-834-291-15
Perfect score: 20
Sequence: 1 aatgttcttaagattttc 20

Scoring table: IDENTITY NUC
Gapco 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	645	21	AAA75620
2	18.4	92.0	645	24	AA520735
3	18.4	92.0	2696	23	ABL27256
4	18.4	92.0	3273	23	ABL27256
5	18.4	92.0	4839	23	ABL14642
6	18.4	92.0	6164	24	ABL32039
7	17.4	87.0	149	21	AAC29641
8	17.4	87.0	536	24	ABN63544

C	9	17.4	87.0	954	24	ABN71515
C	10	17.4	87.0	996	24	ABN66487
C	11	17.4	87.0	1068	20	ABN66079
C	12	17.4	87.0	1528	20	AA339670
C	13	17.4	87.0	1528	24	ABK84789
C	14	17.4	87.0	2483	23	ABL16936
C	15	17.4	87.0	2531	23	ABL11262
C	16	17.4	87.0	4064	24	ABK84306
C	17	17.4	87.0	4664	23	ABL12292
C	18	17.4	87.0	5519	24	ABK39970
C	19	17.4	87.0	215561	24	ABN71527
C	20	16.8	84.0	20	21	AA288724
C	21	16.8	84.0	313	24	ABN20279
C	22	16.8	84.0	366	22	AA135113
C	23	16.8	84.0	366	25	ABK58101
C	24	16.8	84.0	401	22	AA192787
C	25	16.8	84.0	405	22	AA187334
C	26	16.8	84.0	438	22	AAK56912
C	27	16.8	84.0	439	22	AAK69090
C	28	16.8	84.0	450	25	ABK47860
C	29	16.8	84.0	473	21	AAK22374
C	30	16.8	84.0	486	24	ABN63569
C	31	16.8	84.0	515	22	AA46922
C	32	16.8	84.0	515	24	ABK81700
C	33	16.8	84.0	590	24	ABN63363
C	34	16.8	84.0	993	23	AAK53595
C	35	16.8	84.0	1462	24	ABK54435
C	36	16.8	84.0	1608	24	ABL34418
C	37	16.8	84.0	1843	24	ABZ15263
C	38	16.8	84.0	2121	22	AAH14223
C	39	16.8	84.0	2742	21	AAH4257
C	40	16.8	84.0	3481	23	AB106966
C	41	16.8	84.0	3562	22	ABK17959
C	42	16.8	84.0	5252	24	ABL31135
C	43	16.8	84.0	5277	23	ABL18542
C	44	16.8	84.0	5816	24	ABK40012
C	45	16.8	84.0	5937	24	ABN80151

ALIGNMENTS

RESULT 1
AAA75620/c
AAA75620 standard; DNA; 645 BP.

22-JAN-2001 (first entry)

Nucleotide sequence of a human zalpha1 ligand fragment.
zalpha1 ligand; cytokine; haematopoietic cell proliferation; lymphoma;
tumorigenesis; leukaemia; hematopoiesis; B cell tumour; ss.

Location/Qualifiers
436..513
/*tag= a
/number= 4
514..603
/*tag= b
/number= 4
604..645
/*tag= c
/number= 5
/note= "partial exon sequence"

W0200053761-A2.
14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06067.
XX
XX 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI; 2000-565600/52.
XX
PT New human cytokine, designated zalphal ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
XX
PS Example 52; Page 251, 256pp; English.
XX
CC The present sequence represents part of the human zalphal ligand gene.
CC zalphal ligand is a cytokine. The zalphal ligand is useful for
CC stimulating the proliferation and development of haematopoietic cells
CC in vitro and in vivo. Zalphal ligand polynucleotides can be used as
CC primers or probes for cloning the zalphal gene. The zalphal ligand is
CC useful for treating tumourigenesis. A zalphal ligand-saporin fusion
CC toxin may be used for treating leukaemias and lymphomas. Antagonists
CC against zalphal ligand are useful as research reagents for
CC characterizing ligand-receptor interaction. Antagonists are also useful
CC for inhibiting ligand expansion, proliferation, activation and differentiation
CC of cells involved in regulating hematopoiesis. The zalphal ligand may
CC also be used to stimulate an immune response against B cell tumour, a
CC virus, a parasite or a bacterium. The zalphal polypeptides,
CC polynucleotides, antagonists, agonists and antibodies are also useful
CC for the detection, diagnosis, prevention, and treatment of diseases
CC associated with a zalphal ligand genetic defect.
XX
SQ Sequence 645 BP; 244 A; 107 C; 89 G; 204 T; 1 other;
XX
Query Match 92.0%; Score 18.4; DB 21; Length 645;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AATGTTCTTAAGATTGTTT 20
DB 383 AATGTTCTTAAGATTGTTT 364

RESULT 2
AAS20735/c
ID AAS20735 standard; DNA; 645 BP.
XX
AC AAS20735;
XX
DT 09-APR-2002 (first entry)
XX
DE Human zalphal ligand gene partial intron 3-partial exon 5 sequence.
XX
XX Cytokine; zalphal ligand; zalphal receptor; NK cell progenitor;
XX natural killer cell proliferation; T-cell proliferation;
XX B-cell proliferation; anti-tumour response; immune system;
XX immunostimulant; cytoactive; human; ds.
XX
OS Homo sapiens.
XX
PN US6307024-B1.
XX
PD 23-OCT-2001.
XX
PF 09-MAR-2000; 2000US-0522217.
XX
PR 09-MAR-1999; 99US-123547P.
PR 11-MAR-1999; 99US-123904P.
PR 01-JUL-1999; 99US-142013P.

XX
PA (ZYMO) ZYMOGENETICS INC.
XX
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI; 2002-040208/05.
XX
XX
PT New zalphal ligand polypeptides and polynucleotides, useful for
PT stimulating proliferation, activation, differentiation and/or induction
PT of inhibition of specialized cell function, or for stimulating an
PT antigenic response -
XX
PS Example 52; Column 187-188; 105pp; English.
XX
CC The present invention relates to the isolation of a novel cytokine,
CC zalphal ligand and the polynucleotide encoding it. The invention
CC also gives the sequence for the zalphal receptor and the polynucleotide
CC encoding it. The zalphal ligand polypeptide stimulates proliferation of
CC natural killer (NK) cells or NK cell progenitors, the activation of NK
CC cells, proliferation of T-cells, proliferation of B-cells stimulated
CC with anti-CD40 antibodies, stimulates an antigenic response in a mammal,
CC and reduces proliferation of B-cells stimulated with anti-1gM antibodies.
CC The zalphal ligand polypeptide is also useful in preparing antibodies
CC that bind to zalphal ligand epitopes. The zalphal ligand
CC polynucleotides can be used as probes or primers to clone regions
CC of a zalphal ligand gene, and in gene therapy. Zalphal ligand may
CC also be used to identify inhibitors of its activity, to enhance the
CC generation of anti-tumour responses with or without the infusion of
CC donor lymphocytes, and to activate or stimulate the immune system.
CC The present sequence represents part of the human zalphal ligand gene
CC containing the partial intron 3, exon 4, intron 4 and partial exon 5
CC sequences.
XX
SQ Sequence 645 BP; 244 A; 107 C; 89 G; 204 T; 1 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 645;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AATGTTCTTAAGATTGTTT 20
DB 383 AATGTTCTTAAGATTGTTT 364

RESULT 3
ABL22726/c
ID ABL22726 standard; DNA; 2696 BP.
XX
AC ABL22726;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 19651.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX
XX 27-SEP-2001.
XX
PD 23-MAR-2001; 2001WO-US09231.
XX
PF 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
PA Venter JC, Adams M, Li PWD, Myers EW;
PI

DR WPI; 2001-656860/75.
 CC New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 19651; 21pp + Sequence Listing; English.
 PS
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins
 CC sequences (AB101840-AB116175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 2696 BP; 919 A; 581 C; 512 G; 684 T; 0 other;
 SQ
 Query Match 92.0%; Score 18.4; DB 23; Length 2696;
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AATGTTCTTAAGATTGTTT 20
 Db 1704 AATGTTCTTAAGATTGTTT 1685
 RESULT 4
 ID ABL27256 standard; DNA; 3273 BP.
 AC ABL27256;
 XX 26-MAR-2002 (first entry)
 DT
 XX *Drosophila melanogaster* genomic polynucleotide SEQ ID NO 33241.
 DE
 XX *Drosophila*; developmental biology; cell signalling; insecticide;
 XX *Drosophila*; gene; ds.
 KM pharmaceutical; gene; ds.
 KM
 XX *Drosophila melanogaster*.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PR
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 33241; 21pp + Sequence Listing; English.
 PS
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB101840-AB116175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 3273 BP; 1076 A; 658 C; 639 G; 900 T; 0 other;
 SQ
 Query Match 92.0%; Score 18.4; DB 23; Length 3273;
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AATGTTCTTAAGATTGTTT 20
 Db 776 AATGTTCTTAAGATTGTTT 795
 RESULT 5
 ID ABL14642/c standard; cDNA; 4839 BP.
 AC ABL14642;
 XX 26-MAR-2002 (first entry)
 DT
 XX *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 38408.
 DE
 XX *Drosophila*; developmental biology; cell signalling; insecticide;
 XX *Drosophila*; gene; ss.
 KM pharmaceutical; gene; ss.
 KM
 XX *Drosophila melanogaster*.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PR
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX P-PSDB; ABB70539.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 38408; 21pp + Sequence Listing; English.
 PS
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC sequences (AB101840-AB116175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 4839 BP; 1360 A; 1051 C; 1092 G; 1336 T; 0 other;
 SQ
 Query Match 92.0%; Score 18.4; DB 23; Length 4839;
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;
 Matches 19; Conservative 1; Indels 0; Gaps 0;
 QY 1 AATGTTCTTAAGATTGTTT 20

DB 2454 AATGTTCTTAAGATTITTT 2435

|||||

RESULT 6
ID ABL32039 standard; DNA; 6164 BP.
XX
AC ABL32039;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 12.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anti-anemic; cytosine; noctropic;
KW antiproteolytic; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antirheumatic; antidiabetic; antiparasitic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
FN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation.
XX
PS Claim 1; SEQ ID NO 12; 32pp + Sequence listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6164 BP; 1716 A; 134 C; 1458 G; 2855 T; 1 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 6164;
Best Local Similarity 95.0%; Pred. No. 2.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 AATGTTCTTAAGATTITTT 20
DB 4643 AATGTTCTTAAGATTITTT 4662

RESULT 7
ID AAC29641 standard; CDNA; 149 BP.
XX
AC AAC29641;
XX
DT 06-OCT-2000 (first entry)

XX
DE Human secreted protein 5' EST, SEQ ID NO: 33716.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
FN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GIST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 33716; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC cDNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 149 BP; 39 A; 24 C; 25 G; 61 T; 0 other;
XX
Query Match 87.0%; Score 17.4; DB 21; Length 149;
Best Local Similarity 94.7%; Pred. No. 6.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 AATGTTCTTAAGATTITTT 19
DB 37 AATGTTCTTAAGATTITTT 55

RESULT 8
ID ABN63544/c standard; CDNA; 556 BP.
XX
AC ABN63544;
XX
DT 28-JUN-2002 (first entry)
XX
DE Human cancer related polynucleotide SEQ ID NO 3511.
XX
KW Human; cytosine; gene expression; gene mapping; tissue profiling;
KW gene therapy; cancer; tumour; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200214500-A2.
XX
PD 21-FEB-2002.

PF 16-AUG-2001; 2001WO-US25840.
XX
PR 16-AUG-2000; 2000US-226326P.
XX
PA (CHIR-) CHIRON CORP.
XX (HXS-) HXSQ INC.
XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F,
PI Lamson G, Scott EM, Zhang G, Kaessam A, Pot D, Labat I,
XX
DR MPI; 2002-241905/29.
XX
PT New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth -
XX
PS Claim 1; SEQ ID NO 3511; 883bp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytosolic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumor growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 556 BP; 186 A; 111 C; 120 G; 139 T; 0 other;
XX
Query Match 87.0%; Score 17.4; DB 24; Length 556;
Best Local Similarity 94.7%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 ATGTTCTTAAGATTCTTT 20
DB 245 ATGTTCTTAAGATTCTTT 227
XX
RESULT 9
ABN71515/c
ID ABN71515 standard; DNA; 954 BP.
XX
AC ABN71515;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 10943.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; de.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Frazer C,
PI Tectelin H;
XX

DR MPI; 2002-352536/38.
XX
DR P-PSDB; ABP30884.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 7; Page 4191; 4525bp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 954 BP; 293 A; 145 C; 217 G; 299 T; 0 other;
XX
Query Match 87.0%; Score 17.4; DB 24; Length 954;
Best Local Similarity 94.7%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 ATGTTCTTAAGATTCTTT 20
DB 764 ATGTTCTTAAGATTCTTT 746
XX
RESULT 10
ABN66487/c
ID ABN66487 standard; DNA; 996 BP.
XX
AC ABN66487;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 887.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; de.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Frazer C,
PI Tectelin H;
XX
DR MPI; 2002-352536/38.
DR P-PSDB; ABP25856.
XX

PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -

PS Claim 7; Page 3241-3242; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.

SQ Sequence 996 BP; 309 A; 148 C; 229 G; 310 T; 0 other;

Query Match 87.0%; Score 17.4; DB 24; Length 996;
 Best Local Similarity 94.7%; Pred. No. 6.5e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTCTTAAGATTTT 20
 ||| ||||| ||||| |||||

Db 806 ATGTTCTTAAGATTTT 788

RESULT 11

ABN66079/C
 ID ABN66079 standard; DNA; 1068 BP.

AC ABN66079;

DT 01-JUL-2002 (first entry)

DE Streptococcus polymucoside SEQ ID NO 71.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KM group A streptococcus; Streptococcus pyogenes; antibacterial; gene;

OS antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

XX Streptococcus pyogenes.

XX WO200234771-A2.

PD 02-MAY-2002.

PR 29-OCT-2001; 2001WO-GB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.

PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tettelein H;

DR WPI; 2002-352536/38.

DR P-PSDB; ABP25448.

XX New Streptococcus protein for the treatment or prevention of infection

PT or disease caused by Streptococcus bacteria, such as meningitis, and

PT for detecting a compound that binds to the protein -

XX Claim 7; Page 3161; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.

SQ Sequence 1068 BP; 341 A; 204 C; 224 G; 299 T; 0 other;

Query Match 87.0%; Score 17.4; DB 24; Length 1068;
 Best Local Similarity 94.7%; Pred. No. 6.4e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTCTTAAGATTTT 20
 ||| ||||| ||||| |||||

Db 1043 ATGTTCTTAAGATTTT 1025

RESULT 12

AAX39670
 ID AAX39670 standard; DNA; 1528 BP.

AC AAX39670;

DT 02-JUL-1999 (first entry)

DE Renal cancer associated gene.

XX Cancer associated antigen; diagnosis; research; treatment; human;

KM breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;

OS prostate cancer; ss.

XX Homo sapiens.

PN WO9904265-A2.

XX 28-JAN-1999.

PR 15-JUL-1998; 98WO-US14679.

PR 22-JUN-1998; 98US-0102322.

PR 17-JUL-1997; 97US-0896164.

PR 10-OCT-1997; 97US-0061599.

PR 10-OCT-1997; 97US-0061765.

PR 10-OCT-1997; 97US-0948705.

PR 11-OCT-1997; 97GB-0021697.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;

PI Pfeundschuh M, Sahin U, Scanlan MJ, Stockert E;

PI Tureci O;

DR WPI; 1999-132448/11.

XX New isolated cancer associated nucleic acids and polypeptides -

PT isolated using sera from cancer patients, used to develop products

PT for the diagnosis, monitoring or treatment of cancers

DR WPI; 2001-656860/75.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1; SEQ ID NO 1981; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 2483 BP; 787 A; 474 C; 454 G; 768 T; 0 other;
 OY
 Query Match 87.0%; Score 17.4; DB 23; Length 2483;
 Best Local Similarity 94.7%; Pred. No. 6.2e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 2 AATGTTCTTAAGATTTT 20
 2339 AAGTTCTTAAGATTTT 2357
 DB
 RESULT 15
 ABL1262
 ID ABL1262 standard; cDNA; 2531 BP.
 XX
 AC ABL1262;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 28268.
 XX
 KW *Drosophila*; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS *Drosophila melanogaster*.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 PT WPI; 2001-656860/75.
 DR P-PSDB; ABB67159.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1; SEQ ID NO 28268; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 2531 BP; 748 A; 489 C; 587 G; 707 T; 0 other;
 OY
 Query Match 87.0%; Score 17.4; DB 23; Length 2531;
 Best Local Similarity 94.7%; Pred. No. 6.2e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AATGTTCTTAAGATTTT 19
 2253 AATGTTCTTAAGATTTT 2271
 DB
 Search completed: August 1, 2003, 13:55:23
 Job time : 8.69868 secs

Sun Aug 3 09:04:02 2003

US-09-834-291-15.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds
(without alignments)
7955.924 Million cell updates/sec

Title: US-09-834-291-15

Perfect score: 1 aatgcttctaagattttt 20

Sequence: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Scoring table: 1439767 seqs, 1031500376 residues

Searched: 2879534

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	20	100.0	20	9	US-09-834-291-15	Sequence 15, Appl
2	18.4	92.0	645	10	US-09-923-246-108	Sequence 108, App
3	18.4	92.0	645	14	US-10-925-723-108	Sequence 108, App
4	18	90.0	587	13	US-10-027-632-224463	Sequence 224463, A
5	17.4	87.0	733	13	US-10-027-632-21892	Sequence 21892, A
6	16.8	84.0	20	9	US-09-834-291-14	Sequence 14, Appl
7	16.8	84.0	40	9	US-09-834-291-20	Sequence 20, Appl
8	16.8	84.0	40	9	US-09-834-291-28	Sequence 28, Appl
9	16.8	84.0	366	10	US-09-764-877-455	Sequence 455, App
10	16.8	84.0	450	10	US-09-960-352-13025	Sequence 13025, A
11	16.8	84.0	479	11	US-09-918-995-21175	Sequence 21175, A
12	16.8	84.0	515	9	US-09-811-284-104	Sequence 104, App
13	16.8	84.0	578	13	US-10-027-632-245649	Sequence 245649, A
14	16.8	84.0	583	13	US-10-027-632-229994	Sequence 229994, A
15	16.8	84.0	603	13	US-10-027-632-239669	Sequence 239669, A
16	16.8	84.0	650	13	US-10-027-632-200121	Sequence 200121, A

17	16.8	84.0	666	13	US-10-027-632-256867	Sequence 256867, A
18	16.8	84.0	666	13	US-10-027-632-256868	Sequence 256868, A
19	16.8	84.0	666	13	US-10-027-632-256869	Sequence 256869, A
20	16.8	84.0	666	13	US-10-027-632-256870	Sequence 256870, A
21	16.8	84.0	666	13	US-10-027-632-256871	Sequence 256871, A
22	16.8	84.0	666	13	US-10-027-632-170398	Sequence 170398, A
23	16.8	84.0	925	13	US-10-027-632-121985	Sequence 121985, A
24	16.8	84.0	980	13	US-10-027-632-9152	Sequence 9152, App
25	16.8	84.0	993	9	US-09-815-242-7222	Sequence 7222, App
26	16.8	84.0	1224	13	US-10-027-632-203488	Sequence 203488, A
27	16.8	84.0	1843	10	US-09-938-842A-3068	Sequence 3068, App
28	16.8	84.0	2272	14	US-10-114-170-126	Sequence 126, App
29	16.8	84.0	2827	9	US-09-834-291-4	Sequence 4, Appl
30	16.8	84.0	3212	9	US-09-834-291-1	Sequence 1, Appl
31	16.8	84.0	7525	10	US-09-967-552A-72	Sequence 72, Appl
32	16.8	84.0	27483	10	US-09-764-877-2928	Sequence 2928, App
33	16.8	84.0	640681	10	US-09-790-988-1	Sequence 1, Appl
34	16.8	84.0	169139	14	US-10-067-514-1	Sequence 1, Appl
35	16.4	82.0	336	10	US-09-834-975-358	Sequence 358, App
36	16.4	82.0	432	11	US-09-918-995-17530	Sequence 17530, A
37	16.4	82.0	510	10	US-09-833-381-1375	Sequence 1375, App
38	16.4	82.0	549	14	US-10-066-543-2878	Sequence 2878, App
39	16.4	82.0	571	13	US-10-027-632-193079	Sequence 193079, A
40	16.4	82.0	602	13	US-10-027-632-13561	Sequence 13561, A
41	16.4	82.0	602	13	US-10-027-632-13562	Sequence 13562, A
42	16.4	82.0	617	13	US-10-027-632-304588	Sequence 304588, A
43	16.4	82.0	660	13	US-10-027-632-255212	Sequence 255212, A
44	16.4	82.0	684	13	US-10-027-632-84724	Sequence 84724, A
45	16.4	82.0	2000	10	US-09-938-842A-4583	Sequence 4583, App

ALIGNMENTS

RESULT 1
US-09-834-291-15
; Sequence 15, Application US/09834291
; Patent No. US2002042064A1
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; APPLICANT: Muller-Schilling, Martina
; APPLICANT: Oren, Moshe
; TITLE OF INVENTION: p53 Binding Areas
; FILE REFERENCE: 4121-122
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/DE99/03343
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: DE 198 47 779.1
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-834-291-15
Query Match 100.0%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 AATGTTCTTAAGATTTTT 20
Db 1 AATGTTCTTAAGATTTTT 20
RESULT 2
US-09-923-246-108/c
; Sequence 108, Application US/09923246
; Patent No. US20020128446A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020128446A1ak, Julia E.

```
APPLICANT: Preenell, Scott R.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Foster, Donald C.
APPLICANT: Holly, Richard D.
APPLICANT: Gross, Jane A.
APPLICANT: Johnston, Janet V.
APPLICANT: Nelson, Andrew J.
APPLICANT: Dillon, Stacey R.
APPLICANT: Hammond, Angela K.
TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
FILE REFERENCE: 99-16
CURRENT APPLICATION NUMBER: US/09/923,246
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/522,217
PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/123,904
PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/142,013
NUMBER OF SEQ ID NOS: 115
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 108
LENGTH: 645
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(645)
OTHER INFORMATION: n = A,T,C or G
US-09-923-246-108
```

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 10; Length 645;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 AATGTTCTTAAGATTTT 20
Db 383 AATGTTCTTAAGATTTT 364
```

```
RESULT 3
US-10-295-723-108/c
Sequence 108, Application US/10295723
Publication No. US20030125524A1
GENERAL INFORMATION:
APPLICANT: No. US20030125524A1ak, Julia E.
APPLICANT: Preenell, Scott R.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Foster, Donald C.
APPLICANT: Holly, Richard D.
APPLICANT: Gross, Jane A.
APPLICANT: Johnston, Janet V.
APPLICANT: Nelson, Andrew J.
APPLICANT: Dillon, Stacey R.
APPLICANT: Hammond, Angela K.
TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
FILE REFERENCE: 99-16
CURRENT APPLICATION NUMBER: US/10/295,723
PRIOR FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: 09/522,217
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: US 60/123,547
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: US 60/123,904
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: US 60/142,013
NUMBER OF SEQ ID NOS: 115
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 108
LENGTH: 645
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(645)
OTHER INFORMATION: n = A,T,C or G
US-10-295-723-108
```

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 14; Length 645;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 AATGTTCTTAAGATTTT 20
Db 383 AATGTTCTTAAGATTTT 364
```

```
RESULT 4
US-10-027-632-224463
Sequence 224463, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 224463
LENGTH: 587
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(587)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-224463
```

Query Match

Best Local Similarity 100.0%; Score 18; DB 13; Length 587;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 2 AATGTTCTTAAGATTTT 19
Db 134 AATGTTCTTAAGATTTT 151
```

```
RESULT 5
US-10-027-632-21892/c
Sequence 21892, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
```

PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21892
LENGTH: 733
TYPE: DNA
ORGANISM: Human
US-10-027-632-21892

Query Match 87.0%; Score 17.4; DB 13;
Best Local Similarity 94.7%; Pred. No. 7.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19
DB 253 AATGTTCTTAAGATTTT 235

RESULT 6
US-09-834-291-14
Sequence 14, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-14

Query Match 84.0%; Score 16.8; DB 9;
Best Local Similarity 90.0%; Pred. No. 8.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
DB 1 AATGTTCTTAAGATTTT 20

RESULT 7
US-09-834-291-20
Sequence 20, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21

PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 40
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-20

Query Match 84.0%; Score 16.8; DB 9;
Best Local Similarity 90.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
DB 11 AATGTTCTTAAGATTTT 30

RESULT 8
US-09-834-291-28
Sequence 28, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 40
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-28

Query Match 84.0%; Score 16.8; DB 9;
Best Local Similarity 90.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
DB 11 AATGTTCTTAAGATTTT 30

RESULT 9
US-09-764-877-455/c
Sequence 455, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 455
LENGTH: 366
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE

/ LOCATION: (302)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (310)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (313)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-877-455

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 366;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20
DB 274 AATGTTCTTAAGATTGTTT 255

RESULT 10
US-09-960-352-13025
/ Sequence 13025, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathalagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ NUMBER OF SEQ ID NOS: 2001-09-24
/ SEQ ID NO 13025
/ LENGTH: 450
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 56-BOWMS1-023-Q1-E1-F8
US-09-960-352-13025

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 450;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20
DB 127 AATGTTCTTAAGATTGTTT 146

RESULT 11
US-09-918-995-21175/C
/ Sequence 21175, Application US/09918995
/ Publication No. US20030073623A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/ FILE REFERENCE: 20411-756
/ CURRENT APPLICATION NUMBER: US/09/918,995
/ PRIOR FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: US/09/235,076
/ NUMBER OF SEQ ID NOS: 1999-01-20
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 21175
/ LENGTH: 479
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(479)
/ OTHER INFORMATION: n = A,T,C or G

US-09-918-995-21175

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 11; Length 479;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20
DB 152 AATGTTCTTAAGATTGTTT 133

RESULT 12
US-09-811-284-104/C
/ Sequence 104, Application US/09811284
/ Patent No. US20020058306A1
/ GENERAL INFORMATION:
/ APPLICANT: Vogelf, Gabriel
/ FILE REFERENCE: 00167US1
/ CURRENT APPLICATION NUMBER: US/09/811,284
/ PRIOR FILING DATE: 2001-03-16
/ PRIOR APPLICATION NUMBER: 60/189,783
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/189,907
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/189,918
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/189,960
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/189,917
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/192,945
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/192,916
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/192,923
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/192,933
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/192,830
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/192,234
/ PRIOR FILING DATE: 2000-03-27
/ PRIOR APPLICATION NUMBER: 60/192,155
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,935
/ PRIOR FILING DATE: 2000-03-29
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 104
/ LENGTH: 515
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-811-284-104

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 9; Length 515;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20
DB 299 AATGTTCTTAAGATTGTTT 280

RESULT 13
US-10-027-632-245649/C
/ Sequence 245649, Application US/10027632
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 245649
 ; LENGTH: 578
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-245649

Query Match 84.0%; Score 16.8; DB 13; Length 578;
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
 DB 499 AATGTTCTTAAGATTTT 480

RESULT 14
 US-10-027-632-229994
 ; Sequence 229994, Application US/10027632
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 229994
 ; LENGTH: 583
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-229994

Query Match 84.0%; Score 16.8; DB 13; Length 583;
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
 DB 123 AATGTTCTTAAGATTTT 142

RESULT 15
 US-10-027-632-239669/c
 ; Sequence 239669, Application US/10027632
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 239669
 ; LENGTH: 603
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(603)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-027-632-239669

Query Match 84.0%; Score 16.8; DB 13; Length 603;
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20
 DB 262 AATGTTCTTAAGATTTT 243

Search completed: August 1, 2003, 13:37:02
 Job time : 6.18607 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.959 Million cell updates/sec

Title: US-09-834-291-15

Perfect score: 20
Sequence: 1 aatgttcttaagattttt 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCITUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	645	4	US-09-522-217-108
C 2	16.8	84.0	2272	4	US-09-453-702B-116
C 3	16.4	82.0	1770	3	US-08-943-731-146
C 4	16.4	82.0	2388	4	US-09-205-258-154
C 5	16.4	82.0	24183	3	US-08-943-731-3
C 6	16.4	82.0	1664976	4	US-08-916-421B-1
C 7	15.8	79.0	672	4	US-09-134-001C-1361
C 8	15.8	79.0	1619	1	US-08-507-455-1
C 9	15.8	79.0	1644	4	US-08-759-436-4
C 10	15.8	79.0	1554	1	US-08-507-455-2
C 11	15.8	79.0	3084	4	US-09-620-312D-616
C 12	15.8	79.0	48974	3	US-08-920-422-17
C 13	15.8	79.0	580073	4	US-08-545-528D-1
C 14	15.4	77.0	339	4	US-09-107-532A-3414
C 15	15.4	77.0	1020	4	US-09-107-532A-1250
C 16	15.4	77.0	2993	2	US-08-415-593-42
C 17	15.4	77.0	6718	2	US-08-962-284-1
C 18	15.4	77.0	7042	4	US-09-092-508-1
C 19	15.4	77.0	7042	4	US-09-435-115-1
C 20	15.4	77.0	7042	4	US-09-098-310-1
C 21	15.4	77.0	7042	4	US-09-690-364-21
C 22	15.4	77.0	7075	3	US-09-092-508-15
C 23	15.4	77.0	7075	3	US-09-435-115-15
C 24	15.2	76.0	106	2	US-08-184-009-165
C 25	15.2	76.0	106	2	US-08-458-356-165
C 26	15.2	76.0	106	2	US-08-460-736-165
C 27	15.2	76.0	106	4	US-09-535-370-165

C 28	15.2	76.0	264	2	US-08-607-384A-7	Sequence 7, Appl1
C 29	15.2	76.0	363	4	US-09-107-532A-1913	Sequence 1913, Ap
C 30	15.2	76.0	455	6	5278286-1	Patent No. 5278286
C 31	15.2	76.0	457	4	US-09-702-705-861	Sequence 861, App
C 32	15.2	76.0	457	4	US-09-736-457-861	Sequence 861, App
C 33	15.2	76.0	528	4	US-09-702-705-957	Sequence 957, App
C 34	15.2	76.0	528	4	US-09-736-457-957	Sequence 957, App
C 35	15.2	76.0	541	4	US-09-679-409-37	Sequence 37, Appl1
C 36	15.2	76.0	601	2	US-08-184-009-168	Sequence 168, App
C 37	15.2	76.0	601	2	US-08-458-356-168	Sequence 168, App
C 38	15.2	76.0	601	3	US-08-460-736-168	Sequence 168, App
C 39	15.2	76.0	601	4	US-09-535-370-168	Sequence 168, App
C 40	15.2	76.0	687	4	US-09-328-352-1113	Sequence 1113, Ap
C 41	15.2	76.0	705	4	US-09-107-532A-2425	Sequence 2425, Ap
C 42	15.2	76.0	789	4	US-09-134-001C-2695	Sequence 2695, Ap
C 43	15.2	76.0	813	3	US-09-028-819-13	Sequence 13, Appl1
C 44	15.2	76.0	813	4	US-09-556-163-13	Sequence 13, Appl1
C 45	15.2	76.0	915	4	US-09-601-198-90	Sequence 90, Appl1

ALIGNMENTS

RESULT 1

US-09-522-217-108/c
Sequence 108, Application US/09522217
Patent No. 6307024

GENERAL INFORMATION:

APPLICANT: NO. 6307024ak, Julia B.

APPLICANT: Presnell, Scott R.

APPLICANT: Sprecher, Cindy A.

APPLICANT: Foster, Donald C.

APPLICANT: Holly, Richard D.

APPLICANT: Gross, Jane A.

APPLICANT: Johnston, Janet V.

APPLICANT: Nelson, Andrew J.

APPLICANT: Dillon, Stacey R.

APPLICANT: Hammond, Angela R.

TITLE OF INVENTION: NOVEL CYTOKINE ZALPHAI1 LIGAND

FILE REFERENCE: 99-16

CURRENT APPLICATION NUMBER: US/09/522,217

EARLIER FILING DATE: 2000-03-09

EARLIER APPLICATION NUMBER: US 60/123,547

EARLIER FILING DATE: 1999-03-09

EARLIER APPLICATION NUMBER: US 60/123,904

EARLIER FILING DATE: 1999-03-11

EARLIER APPLICATION NUMBER: US 60/142,013

NUMBER OF SEQ ID NOS: 115

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 108

LENGTH: 645

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(645)

OTHER INFORMATION: n = A,T,C or G

US-09-522-217-108

Query Match 92.0%; Score 18.4; DB 4; Length 645;

Best Local Similarity 95.0%; Pred. No. 39;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTC 20

DB 383 AATGTTCTTAAGATTTC 364

RESULT 2

US-09-453-702B-126/c

Sequence 126, Application US/09453702B

Patent No. 6365723

GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6265723e1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Discrete, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 126:
SEQUENCE CHARACTERISTICS:
LENGTH: 2272
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 126:
US-09-453-702B-126
Query Match 84.0%; Score 16.8; DB 4; Length 2272;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AATGTTCTTAAGATTTT 20
DB 1176 AATGTGCTTATGATTTT 1157
RESULT 3
US-08-943-731-146
; Sequence 146, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.

STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR.
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 146:
SEQUENCE CHARACTERISTICS:
LENGTH: 1770 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-146
Query Match 82.0%; Score 16.4; DB 3; Length 1770;
Best Local Similarity 94.4%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TGTTCCTTAAGATTTT 20
DB 44 TGTTCCTTAAGATTTAT 61
RESULT 4
US-09-205-258-154
; Sequence 154, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 154
LENGTH: 2388
TYPE: DNA
ORGANISM: Homo sapiens
US-09-205-258-154
Query Match 82.0%; Score 16.4; DB 4; length 2388;

Best Local Similarity 94.4%; Pred. No. 2.1e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;
QY 3 TGTTCCTTAAGATTTT 20
Db 2334 TGTTCCTTAAGATTTT 2351
RESULT 5
US-08-943-731-3
Sequence 3, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTA, CONSTANTINOS D.
APPLICANT: SEREDA, LARISA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORKKO, JARMO
APPLICANT: ALA-KORKKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR.
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA: US 07/803,628
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24183 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-3
Query Match 82.0%; Score 16.4; DB 3; length 24183;
Best Local Similarity 94.4%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;
QY 3 TGTTCCTTAAGATTTT 20
Db 20352 TGTTCCTTAAGATTTT 20369

RESULT 6
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Built et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
PRIORITY FILING DATE: 1997-08-22
PRIORITY FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 82.0%; Score 16.4; DB 4; Length 1664976;
Best Local Similarity 94.4%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATGTTTCTTAAGATTTT 19
Db 387346 ATGTTTCTTAAGATTTT 387329

RESULT 7
US-09-134-001C-1361
Sequence 1361, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1361
LENGTH: 672
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1361

Query Match 79.0%; Score 15.8; DB 4; Length 672;
Best Local Similarity 89.5%; Pred. No. 3.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AATGTTCTTAAGATTTT 19
Db 647 AATGTTCTTAAGATTTT 665

RESULT 8
US-08-507-455-1
Sequence 1, Application US/08507455
Patent No. 5695961
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/507,455
FILING DATE: 08-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9303988.1
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: MITCHARD, LEONARD C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 1498-72
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1619 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 546..547
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 635..636
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 1035..1036
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 1411..1412
FEATURE:
NAME/KEY: misc_feature
LOCATION: 550..555
FEATURE:
NAME/KEY: misc_feature
LOCATION: 574..579
FEATURE:
NAME/KEY: misc_feature
LOCATION: 668..673
FEATURE:
NAME/KEY: misc_feature
LOCATION: 692..697
US-08-507-455-1

Query Match 79.0%; Score 15.8; DB 1; Length 1619;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTCTTAAGATTTT 19
Db 1265 AATGATTATTAAGATTTT 1283

RESULT 9

US-08-759-436-4/c
Sequence 4, Application US/08759436
Patent No. 6437218
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA fragment encoding a protein
TITLE OF INVENTION: Involved in
TITLE OF INVENTION: fatty aldehyde decarboxylase activity, recombinant
TITLE OF INVENTION: molecules comprising said fragment and a method for
TITLE OF INVENTION: obtaining transformed bacterial cells and plants
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Iadec & Parry
STREET: 26 West 61st Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,436
FILING DATE: 5-Dec-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cord, Janet I.
REGISTRATION NUMBER: 33,778
REFERENCE/DOCKET NUMBER: U011063-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-708-1935
TELEFAX: 212-246-8959
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
STRAIN: Landsberg erecta ecotype
IMMEDIATE SOURCE:
LIBRARY: Coates
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 1
MAP POSITION: 12
UNITS: cm
PUBLICATION INFORMATION:
AUTHORS: Aarts, Mark G.M.
AUTHORS: Keijzer, Christiaan J.
AUTHORS: Stiekema, Willem J.
AUTHORS: Pereira, Andy
TITLE: Molecular characterization of the CR1 gene
TITLE: of Arabidopsis involved in epicuticular wax
TITLE: biosynthesis and pollen fertility
JOURNAL: Plant Cell
VOLUME: 7
ISSUE: 12
PAGES: 2115-2127
DATE: december-1995
RELEVANT RESIDUES IN SEQ ID NO: 4: FROM 1 TO 750
US-08-759-436-4

Query Match 79.0%; Score 15.8; DB 4; Length 1644;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTCTTAAGATTTT 19
Db 855 ACTGTTCTGAAGATTTT 837

RESULT 10

US-08-507-455-2
Sequence 2, Application US/08507455
Patent No. 5695961
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/507,455
FILING DATE: 08-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9303988.1
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: MITCHARD, LEONARD C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 1498-72
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1754 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 546..547
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 635..636
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 1035..1036
FEATURE:
NAME/KEY: misc_recomb
LOCATION: 1411..1412
US-08-507-455-2

Query Match 79.0%; Score 15.8; DB 1; Length 1754;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTCTTAAGATTTT 19
Db 1397 AATGATTATTAAGATTTT 1415

RESULT 11

US-09-620-312D-616
Sequence 616, Application US/09620312D
Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Xu, Aigong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc FL_genes Version 1.0
SEQ ID NO 616
LENGTH: 3084
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (633)..(1883)
US-09-620-312D-616

Query Match 79.0%; Score 15.8; DB 4; Length 3084;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19
DB 3012 AAGTTCTTAAGATTTT 3030

RESULT 12

US-08-920-422-17
Sequence 17, Application US/08920422A
Patent No. 6255473

GENERAL INFORMATION:

APPLICANT: Vitek, Michael P.
APPLICANT: Mltuda, No. 6255473iaki
APPLICANT: Roses, Allen D.
TITLE OF INVENTION: Presentin-1 Gene Promoter
FILE REFERENCE: VITEKPRESENTIN
CURRENT APPLICATION NUMBER: US/08/920,422A
CURRENT FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 17
LENGTH: 48974
TYPE: DNA
ORGANISM: Mus musculus
US-08-920-422-17

Query Match 79.0%; Score 15.8; DB 3; Length 48974;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19
DB 15475 AATTTTATTAAGATTTT 15493

RESULT 13

US-08-545-528D-1
Sequence 1, Application US/08545528D
Patent No. 6537773

GENERAL INFORMATION:

APPLICANT: Fraser et al.
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragments
PATENT NO. 6537773
FILE REFERENCE: PB193P1
CURRENT APPLICATION NUMBER: US/08/545,528D
CURRENT FILING DATE: 1995-10-19
PRIOR APPLICATION NUMBER: US 08/488,018
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/473,545
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 580073
TYPE: DNA
ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Query Match 79.0%; Score 15.8; DB 4; Length 580073;
Best Local Similarity 89.5%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19
DB 547054 AATGTTCTTAAGATTTT 547072

RESULT 14

US-09-107-532A-3414
Sequence 3414, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Walham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinifello, Pamela Deneke

REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3414:
SEQUENCE CHARACTERISTICS:
LENGTH: 339 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...339
SEQUENCE DESCRIPTION: SEQ ID NO: 3414:
US-09-107-532A-3414

Query Match 77.0%; Score 15.4; DB 4; Length 339;
Best Local Similarity 94.1%; Pred. No. 5.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAGATT 17
|||||
Db 321 AATGTTCTTAGATT 337

RESULT 15
US-09-107-532A-1250/C
Sequence 1250, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1250:
SEQUENCE CHARACTERISTICS:
LENGTH: 1020 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1020
SEQUENCE DESCRIPTION: SEQ ID NO: 1250:
US-09-107-532A-1250

Query Match 77.0%; Score 15.4; DB 4; Length 1020;
Best Local Similarity 94.1%; Pred. No. 5.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAGATT 17
|||||
Db 201 AATGTTCTTAGATT 185

Search completed: August 1, 2003, 08:37:20
Job time: 4.65066 secs

JOURNAL

Patent: DE 19847779-C 5 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

FEATURES
SOURCE
1. .20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT
ORIGIN
7 a 7 c 5 g 1 t

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGACAGCCCTGACAGCCA 20
1 GGACAGCCCTGACAGCCA 20

RESULT 2

AX026098
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Sequence 10 from Patent DE19847779.
AX026098.1 GI:10187529

Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 10 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

Location/Qualifiers
1. .20

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT
ORIGIN
7 a 7 c 5 g 1 t

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGACAGCCCTGACAGCCA 20
1 GGACAGCCCTGACAGCCA 20

RESULT 3

AX026112
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Sequence 24 from Patent DE19847779.
AX026112.1 GI:10187543

Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 24 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

Location/Qualifiers
1. .20

/organism="Homo sapiens"

BASE COUNT
ORIGIN
7 a 7 c 5 g 1 t

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGACAGCCCTGACAGCCA 20
1 GGACAGCCCTGACAGCCA 20

RESULT 4

AX026120
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Sequence 32 from Patent DE19847779.
AX026120.1 GI:10187551

Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 32 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)

Location/Qualifiers
1. .266

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT
ORIGIN
49 a 72 c 110 g 35 t

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 266;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGACAGCCCTGACAGCCA 20
160 GGACAGCCCTGACAGCCA 179

RESULT 5

HS011034
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens DNA for enhancer of CD95 gene, partial.
HS011034.1 GI:4165483

Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

1 Mueller, M., Wilder, S., Bannasch, D., Isreal, D., Lehlbach, K.,
Li-Weber, M., Friedmann, S.L., Galle, P.R., Stremmel, W., Oren, M. and
Kramer, P.H.
p53 activates the CD95 (APO-1/Fas) gene in response to DNA damage
by anticancer drugs
J. Exp. Med. 188 (11), 2033-2045 (1998)

9841917
2 (bases 1 to 266)

Direct Submission
Mueller, M.

Submitted (08-SEP-1998)

IV, University Hospital, Bergheimerstr. 58, Heidelberg, 69115,

GERMANY
Location/Qualifiers
1. .266
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

gene
/gene="CD95"
159. .178
/gene="CD95"
/evidence="experimental"

enhancer
49 a 72 c 110 g 35 t

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 266;
Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
|||||
Db 160 GGACAGCCCTGACAGCCA 179

RESULT 6
LOCUS AX026090 720 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 2 from Patent DE19847779.
ACCESSION AX026090
VERSION AX026090.1 GI:10187521
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 2 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source Location/Qualifiers
1. .720
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 142 a 181 c 216 g 181 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 720;
Pred. No. 33;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
|||||
Db 160 GGACAGCCCTGACAGCCA 179

RESULT 7
LOCUS AX026091 2380 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 3 from Patent DE19847779.
ACCESSION AX026091
VERSION AX026091.1 GI:10187522
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 3 03-FEB-2000;

DEUTSCHES KREBSFORSCH (DE)
Location/Qualifiers
1. .2380
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 579 a 595 c 568 g 638 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2380;
Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
|||||
Db 1820 GGACAGCCCTGACAGCCA 1839

RESULT 8
LOCUS AX026092 2827 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 4 from Patent DE19847779.
ACCESSION AX026092
VERSION AX026092.1 GI:10187523
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 4 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source Location/Qualifiers
1. .2827
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2827;
Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
|||||
Db 2267 GGACAGCCCTGACAGCCA 2286

RESULT 9
LOCUS AX026089 3212 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 1 from Patent DE19847779.
ACCESSION AX026089
VERSION AX026089.1 GI:10187520
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy
JOURNAL Patent: DE 19847779-C 1 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source Location/Qualifiers
1. .3212
/organism="Homo sapiens"
/mol_type="genomic DNA"

BASE COUNT 778 a 784 c 809 g 841 t /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3212;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGAGCCCTGACAGCCA 20
Db 2659 GGACAGAGCCCTGACAGCCA 2678

RESULT 10
LOCUS AX695635 45121 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 1262 from Patent WO03008583.
ACCESSION AX695635
VERSION AX695635.1 GI:29418787
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Morris,D.W. and Engelhard,E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1262 30-JUN-2003;
Sagres Discovery (US)
FEATURES
source Location/Qualifiers
1..45121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 45121;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGAGCCCTGACAGCCA 20
Db 10653 GGACAGAGCCCTGACAGCCA 10672

RESULT 11
LOCUS AL157394 187313 bp DNA linear PRI 22-AUG-2001
DEFINITION Human DNA sequence from clone RP11-399019 on chromosome 10,
complete sequence.
ACCESSION AL157394
VERSION AL157394.15 GI:15384622
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 187313)
Blakey,S.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Aug 31, 2001 this sequence version replaced gi:14161146.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SwissProt; Tr: TrEMBL; Wp: WormPeP; Information on the WormPeP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-399019 is from the library RPCT-11.2 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pRACE3.6
This sequence is the entire insert of clone RP11-399019 The true left end of clone RP11-496H23 is at 166408 in this sequence. The true right end of clone RP11-30415 is at 18704 in this sequence.

FEATURES
source Location/Qualifiers
1..187313
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

misc_feature
/note="Sequence confirmed by AC015461 sequenced by WIBR."
100119..100156
/clone="RP11-399019"
/clone_1b="RPCT-11.2"

misc_feature
/note="Sequence confirmed by AC015461 sequenced by WIBR."
100157..100198
/note="Sequence from AC015461 sequenced by WIBR."

misc_feature
/note="Sequence from AC015461 sequenced by WIBR."
105808..105972
/note="Sequence from AC015461 sequenced by WIBR."

misc_feature
/note="Sequence confirmed by AC015461 sequenced by WIBR."
105973..105989
/note="Sequence confirmed by AC015461 sequenced by WIBR."

BASE COUNT 55669 a 36398 c 36888 g 58358 t

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 187313;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGAGCCCTGACAGCCA 20
Db 144628 GGACAGAGCCCTGACAGCCA 144647

RESULT 12
LOCUS AX026094 20 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 6 from Patent DE19847779.
ACCESSION AX026094
VERSION AX026094.1 GI:10187525
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Mueller-Schilling,M., Krammer,P. and Oren,M.
TITLE Novel receptor dna useful for identifying apoptosis-modulating
JOURNAL substances potentially useful for cancer chemotherapy
PATENT: DE 19847779-C 6 03-FEB-2000;
DEUTSCHES KREBSFORSCH (DE)
FEATURES
source Location/Qualifiers
1..20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 8 a 6 c 5 g 1 t

Query Match 92.0%; Score 18.4; DB 6; Length 20;
 Best Local Similarity 95.0%; Pred. No. 4.5e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGACAAAGCCCTGACAAAGCA 20
 1 GGAAAAAGCCCTGACAAAGCA 20

RESULT 13
 ALJ356352 167412 bp DNA linear HTG 21-OCT-2001
 LOCUS Homo sapiens chromosome 1 clone RPS-859H16, ** SEQUENCING IN
 DEFINITION PROGRESS ***, 20 unordered pieces.
 ALJ356352
 ALJ356352.8 GI:16304915
 HTG; HTGS_PHASE1; HTGS_CANCELLED.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 McIay, K.
 Direct Submission
 Submitted (18-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Oct 21, 2001 this sequence version replaced gi:12539659.
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Project Information
 Center project name: dJ859H16
 Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 26% of reads
 Dye-terminator Big Dye; 73% of reads
 Consensus quality: 15949 bases at least Q40
 Consensus quality: 162148 bases at least Q30
 Consensus quality: 163926 bases at least Q20
 Insert size: 165512; sum-of-contigs
 Insert size: 206683; 10.2% error; agarose-fp
 Quality coverage: 4.08x in Q20 bases; sum-of-contigs Quality
 coverage: 3.39x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
 consists of 20 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 12019: contig of 12019 bp in length
 12020 12119: gap of 100 bp in length
 12120 32657: contig of 20538 bp in length
 32658 32757: gap of 100 bp
 32758 36105: contig of 3348 bp in length
 36106 36205: gap of 100 bp
 36206 41069: contig of 4864 bp in length
 41070 41169: gap of 100 bp in length
 41170 51654: contig of 10485 bp in length
 51655 51754: gap of 100 bp
 51755 62397: contig of 10643 bp in length
 62398 73595: gap of 100 bp
 73596 73695: contig of 11098 bp in length
 73696 83277: gap of 100 bp
 83278 83377: contig of 9582 bp in length
 83378 83377: gap of 100 bp

FEATURES
 source
 83378 98845: contig of 15468 bp in length
 98846 98945: gap of 100 bp
 98946 101173: contig of 2228 bp in length
 101174 101273: gap of 100 bp
 101274 103420: contig of 2147 bp in length
 103421 103520: gap of 100 bp
 103521 109720: contig of 6200 bp in length
 109721 109820: gap of 100 bp
 109821 114256: contig of 4436 bp in length
 114257 114356: gap of 100 bp
 114357 119627: contig of 5271 bp in length
 119628 119727: gap of 100 bp
 119728 128350: contig of 8623 bp in length
 128351 128451: gap of 100 bp
 128452 140183: contig of 11733 bp in length
 140184 140284: gap of 100 bp
 140285 147703: contig of 7420 bp in length
 147704 147803: gap of 100 bp
 147804 158100: contig of 10297 bp in length
 158101 158200: gap of 100 bp
 158201 161982: contig of 3782 bp in length
 161983 162082: gap of 100 bp
 162083 167412: contig of 5330 bp in length.

Location/Qualifiers
 1. 167412
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone_RPS="RPS-859H16"
 /clone_11b="RPCT-5"
 1. 12019
 /note="assembly_fragment:00380
 clone end:SP6
 vector_side:left"
 12120. 32657
 /note="assembly_fragment:00114
 fragment_chain:1"
 32758. 36105
 /note="assembly_fragment:00063
 fragment_chain:1"
 36206. 41069
 /note="assembly_fragment:01365
 fragment_chain:1"
 41170. 51654
 /note="assembly_fragment:01782
 fragment_chain:1"
 51735. 62397
 /note="assembly_fragment:01003
 fragment_chain:1"
 62498. 73595
 /note="assembly_fragment:00248
 fragment_chain:2"
 73696. 83277
 /note="assembly_fragment:01125
 fragment_chain:2"
 83378. 98845
 /note="assembly_fragment:01780
 fragment_chain:2"
 98946. 101173
 /note="assembly_fragment:01201
 fragment_chain:2"
 101274. 103420
 /note="assembly_fragment:01704
 fragment_chain:3"
 103521. 109720
 /note="assembly_fragment:00350
 fragment_chain:3"
 109821. 114256
 /note="assembly_fragment:01435
 fragment_chain:3"
 114357. 119627
 /note="assembly_fragment:00291

BASE COUNT	ORIGIN
misc_feature	fragment_chain:4"
misc_feature	1197280 .128350
misc_feature	/note="assembly_fragment:01405
misc_feature	fragment_chain:4"
misc_feature	128451 .140183
misc_feature	/note="assembly_fragment:00901
misc_feature	fragment_chain:5"
misc_feature	140284 .147703
misc_feature	/note="assembly_fragment:00215
misc_feature	fragment_chain:5"
misc_feature	147804 .158100
misc_feature	/note="assembly_fragment:00322"
misc_feature	158201 .161982
misc_feature	/note="assembly_fragment:00595"
misc_feature	162083 .167412
misc_feature	/note="assembly_fragment:00872"
misc_feature	41224 c 42341 g 40447 t . 1904 others

Query	1	GGACAGCCTGACAGCCA	20
Query Match	11		Score 18.4; DB 2; Length 167412;
Best Local Similarity	11		Fred. 95.0%; No. 70;
Matches 19; Conservative	11		0; Mismatches 1; Indels 0; Gaps 0;
Db	143124	GGTCAAGCCTGACAGCCA	143105

RESULT 14	
AL672160	
LOCUS	
DEFINITION	AL672160 192657 bp DNA linear ROD 17-APR-2003
ACCESSION	Muscle DNA sequence from clone RP23-144K18 on chromosome 4, complete sequence.
VERSION	AL672160
KEYWORDS	AL672160..10 GI:25809448
SOURCE	HTG.
ORGANISM	Mus musculus (house mouse)
	Mus musculus
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 192657)
REFERENCE	Hammond, S.
AUTHORS	Direct Submission
TITLE	Submitted (17-APR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
JOURNAL	humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Nov 27, 2002 this sequence version replaced gi:24394955.
COMMENT	

Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YVC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

FEATURES	source
Em', EMBL; Sw', SWISSPROT; Tr', TREMBL; Wp', WORMPEEP; Information from the WORMPEEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpeep RP23-144K18 is constructed by the group of Pister de Jong. For further details see http://www.choiri.org/bacpac/home.htm VECTOR: pBACe3.6.	Location/Qualifiers 1..192657 /organism="Mus musculus" /mol_type="Genomic DNA" /db_xref="taxon:10090" /chromosome="4" /clone="RP23-144K18" /clone_id="RP23-23" /clone_id="RP23-23"
BASE COUNT	51940 a 39496 c 42716 g 58505 t
ORIGIN	

	Query Match	92.0%;	Score 18.4;	DB 10;	Length 192657;
	Best Local Similarity	95.0%;	Pred. No. 68;		
Matches 19;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
QY	1 GGACAGCCTGCACAAGCA	20			
Db	156517 GGACAGCCCGACACAAGCA	156536			

RESULT 15				
BX294664/c				
LOCUS	BX294664	193152 bp	DNA	Linear HTG 24-APR-2003
DEFINITION	Mus musculus chromosome X clone RP21-384C22, *** SEQUENCING IN PROGRESS ***, 10 unordered pieces.			
ACCESSION	BX294664			
VERSION	BX294664.2	G1:29500936		
KEYWORDS	HTG; HTGS PHASEI; HTGS CANCELLED.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euarchonta; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 193152)			
	McLay,K.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (23-APR-2003) Wellcome Trust Sanger Institute, Hinxton,			
TITLE	Cambidgeshire, Chr10 ISA, UK. E-mail enquiries:			
JOURNAL	humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk			
COMMENT	On Apr 2, 2003 this sequence version replaced g1:29134752.			

```

from the whole genome shotgun alone has only been used where it has
phred quality of at least 30.

-----Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humgenery@sanger.ac.uk

-----Project Information
Center project name: DM384C22

-----Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 190946 bases at least Q40
Consensus quality: 191521 bases at least Q30
Consensus quality: 191867 bases at least Q20
Insert size: 192452; sum-of-contigs
Insert size: 197663; 4.4% error; agarose-fp
Quality coverage: 8.26x in Q20 bases; sum-of-contigs Quality
coverage: 8.03x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 37461: contig of 37461 bp in length
* 37462 37561: gap of 100 bp
* 37562 45950: contig of 8389 bp in length
* 45951 46051: gap of 100 bp
* 46051 98426: contig of 52376 bp in length
* 98427 98526: gap of 100 bp
* 98527 118027: contig of 19501 bp in length
* 118028 118127: gap of 100 bp
* 118128 126520: contig of 8393 bp in length
* 126521 132811: gap of 100 bp
* 132812 132911: contig of 6191 bp in length
* 132912 172432: gap of 100 bp
* 172433 172532: contig of 39521 bp in length
* 172533 185735: gap of 100 bp
* 185736 185835: contig of 1303 bp in length
* 185836 190693: gap of 100 bp
* 190694 190793: contig of 4858 bp in length
* 190794 193352: gap of 100 bp
* 193352 2559: contig of 2559 bp in length.
Location/Qualifiers
1. 193352

FEATURES

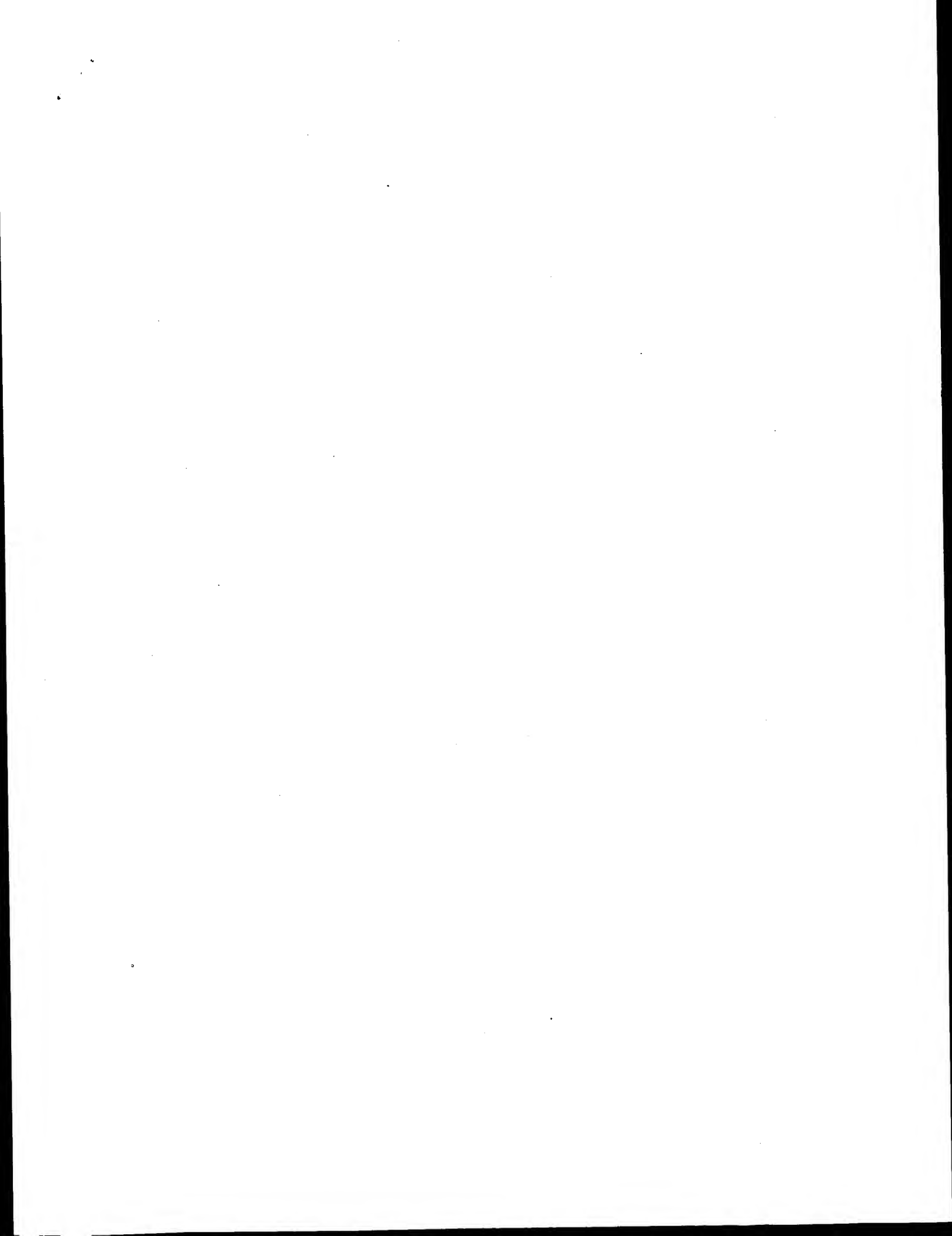
source

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-384C22"
1. 37461
/clone_id="RPCI-23"
/note="assembly fragment:02244
fragment_chain:1
clone_end:SP6
vector_side:left"
37562. 45950
/note="assembly fragment:00826
fragment_chain:1"
46051. 58426
/note="assembly fragment:00297.
fragment_chain:1"
98527. 118027
/note="assembly fragment:01245
fragment_chain:1"
118128. 126520
/note="assembly fragment:02356
fragment_chain:1"
126621. 132811
/note="assembly fragment:01763
fragment_chain:1"
132912. 172432
/note="assembly fragment:01130
fragment_chain:1"
172533. 185735
/note="assembly fragment:01150
fragment_chain:1"
185836. 190693
/note="assembly fragment:00830
fragment_chain:1
clone_end:T7
vector_side:right"
190794. 193352
/note="assembly fragment:00599"
42616 a 42616 c 39284 g 52100 t 900 others

BASE COUNT 58452 a 42616 c 39284 g 52100 t 900 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 193352;
Best Local Similarity 95.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
DB 36135 GGACAGCCCTGACAGCCA 36116



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds
(without alignments) updates/sec
7098.748 Million cell

Title: US-09-834-291-24

Sequence: 1 ggacaagcccgacaagcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_estdb:*
3: em_estdb:*
4: em_estdb:*
5: em_estdb:*
6: em_estdb:*
7: em_estdb:*
8: em_estdb:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estdb:*
16: em_estdb:*
17: em_estdb:*
18: em_estdb:*
19: em_estdb:*
20: em_estdb:*
21: em_estdb:*
22: em_estdb:*
23: em_estdb:*
24: em_estdb:*
25: em_estdb:*
26: em_estdb:*
27: em_estdb:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	793	12	BI763679
2	20	100.0	962	9	AL540709
3	20	100.0	1089	13	BO072834
4	18	90.0	486	14	W45225

Result No.	Score	Query Match	Length	ID	Description
5	18	90.0	506	10	BG659530
6	18	90.0	753	13	BUI42933
7	17.4	87.0	619	28	A2939002
8	17.4	87.0	684	28	A2450876
9	17.4	87.0	689	29	BX220214
10	17.4	87.0	701	14	CA605171
11	17.4	87.0	745	29	BX179961
12	17	85.0	331	14	162311
13	17	85.0	429	28	A2781046
14	16.8	84.0	531	9	AA195416
15	16.8	84.0	592	9	AA131236
16	16.8	84.0	629	28	A2492025
17	16.8	84.0	730	28	BH988275
18	16.8	84.0	766	13	B0471804
19	16.8	84.0	843	14	CB560391
20	16.8	84.0	973	13	B0715536
21	16.8	84.0	1059	10	BF205338
22	16.8	84.0	1077	14	CD519542
23	16.8	84.0	1077	14	AK043145
24	16.8	84.0	1077	14	BF386638
25	16.8	84.0	191	10	BF386638
26	16.8	84.0	262	28	AO632653
27	16.8	84.0	312	12	BI041486
28	16.8	84.0	424	13	BY444129
29	16.8	84.0	545	10	BG376805
30	16.8	84.0	626	28	BH303980
31	16.8	84.0	652	28	AG094812
32	16.8	84.0	675	28	AG094812
33	16.8	84.0	696	10	BF966977
34	16.8	84.0	742	29	B2322825
35	16.8	84.0	802	10	BG481356
36	16.8	84.0	806	10	BG296720
37	16.8	84.0	809	28	AO050636
38	16.8	84.0	890	14	CB192716
39	16.8	84.0	926	13	BH849141
40	16.8	84.0	1114	29	CC278162
41	16.8	84.0	2193	10	BF579152
42	16.8	84.0	300	9	AL916663
43	16.8	84.0	406	14	CB807904
44	16.8	84.0	407	14	CB771883
45	16.8	84.0	427	14	CB795385

ALIGNMENTS

RESULT 1
BI763679/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI763679 793 bp mRNA linear EST 25-SEP-2001
603049567F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189752 5',
mRNA sequence.
BI763679
BI763679.1 GI:15755257
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strussberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1474 row: e column: 17
High quality sequence stop: 786.

FEATURES

source

Location/Qualifiers
1. .793
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5189752"
/lab_host="DH10B"
/clone_1lb="NIH MGC 116"
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC library."

BASE COUNT

ORIGIN

183 a 265 c 208 g 137 t

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 793;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
|||||
Db 398 GGACAAAGCCCTGACAAAGCCA 379

RESULT 2

AL540709/c

LOCUS AL540709 962 bp RNA linear EST 12-MAY-2003
DEFINITION AL540709 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE002YN18
ACCESSION AL540709
VERSION AL540709.2 GI:30544172
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12871113.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5554.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE002DG09P1&cluster=5554.r. Contact :
Peng Liang Email : lliang@lifestech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE002DG09P1.

FEATURES

source

1. .962
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0DE002YN18"
/issue_type="PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 211 a 302 c 264 g 181 t 4 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 962;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
|||||
Db 94 GGACAAAGCCCTGACAAAGCCA 75

RESULT 3

B0072834

LOCUS B0072834 1089 bp RNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6763280 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5756324
ACCESSION B0072834
VERSION B0072834.1 GI:19901880
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/
1 (bases 1 to 1089)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Agencourt Bioscience Corporation (LLNL)
clone distributing by: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2796 row: 1 column: 21
High quality sequence stop: 645.

FEATURES

source

1. .1089
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5756324"
/issue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH MGC library."

BASE COUNT 280 a 263 c 333 g 212 t 1 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 13; Length 1089;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
|||||
Db 369 GGACAAAGCCCTGACAAAGCCA 388

RESULT 4

W45225

LOCUS W45225 486 bp RNA linear EST 10-OCT-1996
DEFINITION zc33c09.r1 Soares_senescent_fibroblasts_NbHSF Homo sapiens CDNA

clone IMAGE:323152 5' similar to gb:125081 TRANSFORMING PROTEIN
 RHOC (HUMAN); mRNA sequence.
 ACCESSION W45225
 VERSION W45225.1 GI:1329306
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 486)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hiltman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
 R., Williamson, A., Woldmann, P. and Wilson, R.
 The WashU-Merck EST Project
 JOURNAL Unpublished
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@image.lbnl.gov) for further information.
 Insert length: 972 Std Error: 0.00
 Seq primer: mob.RBGA+BT
 High quality sequence stop: 296.
 Location/Qualifiers
 FEATURES
 Source
 1. .486
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1254664"
 /db_xref="taxon:9606"
 /clone="IMAGE:323152"
 /issue_type="senescent fibroblast"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pT7T3D (Pharmacia) with a modified
 polylinker V-type phagemid; Site 1: Not I; Site 2: Eco RI
 ; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer [5',
 TGTACCAATCTGAAAGTGGAGCGCGCATTTTCTTTTCTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3D vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 120 a 129 c 140 g 95 t 2 others
 ORIGIN
 Query Match 90.0%; Score 18; DB 14; Length 486;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GACACGCTGACAGCC 19
 DB 265 GACACGCTGACAGCC 282
 RESULT 5
 BG659530 506 bp mRNA linear EST 11-MAY-2001
 LOCUS BG659530
 DEFINITION TgBSTy24901.Y1 TgVEG18 Tachyzoite cDNA library Toxoplasma gondii
 CDNA clone TgBSTy24901.Y1 5' similar to TR:063030 063030 RAT
 ALPHA-SMOOTH MUSCLE ACTIN mRNA FRAGMENT; mRNA sequence.
 ACCESSION BG659530
 VERSION BG659530.1 GI:13801991
 KEYWORDS EST.
 SOURCE Toxoplasma gondii
 ORGANISM Toxoplasma gondii
 BUKARYOTA; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Toxoplasma.

REFERENCE 1 (bases 1 to 506)
 AUTHORS Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,
 Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M.,
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Rletter
 E., Bennett, J., Franklin, C., Tsagaris, R., Ronke, I., Kennedy
 S., Maguire, L., Waterston, R. and Wilson, R.
 Toxoplasma EST Project
 JOURNAL Unpublished
 COMMENT Contact: Clifton, S.
 Toxoplasma EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: tox@wustl.wustl.edu
 Contact David Sibley (tox@est@wustl.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 371.
 Location/Qualifiers
 FEATURES
 Source
 1. .506
 /organism="Toxoplasma gondii"
 /mol_type="mRNA"
 /strain="VEG"
 /db_xref="taxon:5811"
 /clone="TgBSTy24901.Y1"
 /dev_stage="Tachyzoite"
 /lab_host="DH10B"
 /note="Vector: pT7T3D Tachyzoite cDNA library"
 /note="Vector: pT7T3D Tachyzoite cDNA library"
 ; This library was constructed by Kellian Tang, Robert
 Cole, and L. David Sibley at Washington University. CDNA
 were synthesized from poly(A)+ RNA by oligo(dT) priming,
 size-selected and directionally cloned into the uni-zap XR
 lambda vector (Stratagene). The primary library was mass
 excised as phagemids and rescued in SOLR cells. The
 plasmid library was recovered from the SOLR cells and
 transformed in mass into DH10B (GeneHog, Research Genetics
 Inc.) for sequencing. WARNING: This library may contain
 a small percentage contaminants from human fibroblast
 cells."

BASE COUNT 86 a 191 c 140 g 89 t
 ORIGIN
 Query Match 90.0%; Score 18; DB 10; Length 506;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGACAGCCTGACAGC 18
 DB 18 GGACAGCCTGACAGC 1
 RESULT 6
 BU142933/c 753 bp mRNA linear EST 25-NOV-2002
 LOCUS 603135537F1 CSRGCHL25 Gallus gallus cDNA clone CHEST11919 5', mRNA
 DEFINITION
 ACCESSION BU142933
 VERSION BU142933.1 GI:25358894
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 753)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken CDNA
 Curr. Biol. 12 (22), 1965-1969 (2002)
 TITLE JOURNAL
 MEDLINE 22335534
 PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1OD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
Location/Qualifiers
1..753
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHST11919"
/sex="Female"
/tissue_type="cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CEQCH25"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
[stratagene] vector to accommodate cDNA produced with the
T-timed protocol (construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387
,624). Cut pBluescript II KS(+) with NotI and EcoRI
ligate in double stranded adaptor containing BspI and
BamHI sites [5'gagccgctgacccgagtcgcgaataag
(5'aattcttttcggatccgggctgcagc)]"

BASE COUNT 262 a 114 c 152 g 225 t

ORIGIN

Query Match 90.0%; Score 18; DB 13; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCGCTGACAGC 18
|||||

Db 196 GGACAGCGCTGACAGC 179

RESULT 7
LOCUS 619 bp DNA linear GSS 26-APR-2001
DEFINITION 2M0197J18R Mouse 10kb plasmid UUC2M library Mus musculus genomic
clone UUC2M0197J18 R, genomic survey sequence.
ACCESSION A2939002
VERSION A2939002.1 GI:13799261
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 619)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0197 row: J column: 18
Seq primer: CACACAGAAACAGCATGACC

Class: plasmid ends
High quality sequence stop: 619.
Location/Qualifiers
1..619
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC2M0197J18"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUC2M library"
/note="Vector: pMD2env. Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 [gi|4732114|gb|AF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 198 a 116 c 87 g 218 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 619;
Best Local Similarity 94.7%; Pred. No. 7.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGCGCTGACAGC 19
|||||

Db 595 GGACAGCGCTGACAGC 577

RESULT 8
LOCUS 684 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0249M17R Mouse 10kb plasmid UUC2M library Mus musculus genomic
clone UUC2M0249M17 R, genomic survey sequence.
ACCESSION A2450876
VERSION A2450876
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 684)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0249 row: M column: 17

FEATURES

source

Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 684.
 Location/Qualifiers

1..684
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCGCM0249M17"
 /sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUCGCM library"
 /note="Vector: PMD42uv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

180 a 202 c 160 g 142 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 684;
 Best Local Similarity 94.7%; Pred. No. 7.3e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCTGACAGCA 20

DB 611 GACAGCCTGACAGCA 629

RESULT 9

BX220214/c

LOCUS BX220214 689 bp DNA linear GSS 29-JAN-2003
 DEFINITION Danio rerio genomic clone DKEX-269M11, genomic survey sequence.
 ACCESSION BX220214
 VERSION BX220214.1 GI:28052100

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)
 Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 689)
 Humphrey, S.J., Huckle, E. and Durham, J.L.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Unpublished
 This sequence was generated from the SP6 end of BAC 269M11. 269M11 is part of the Daniokey BAC library created by R. Plasterk and N.V. Keygene. Further details:
 http://www.sanger.ac.uk/Projects/D_rerio/
 Location/Qualifiers

FEATURES

source

1..689
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEX-269M11"
 /tissue_type="Testis"

BASE COUNT 180 a 144 c 144 g 221 t
 ORIGIN

Query Match 87.0%; Score 17.4; DB 29; Length 689;
 Best Local Similarity 94.7%; Pred. No. 7.4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCTGACAGCA 20

DB 419 GACAGCCTGACAGCA 401

RESULT 10

CA605171/c

LOCUS CA605171 701 bp mRNA linear EST 21-NOV-2002
 DEFINITION wrl.pk0046.e7 wrl Triticum aestivum cDNA clone wrl.pk0046.e7 5' end
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
 Contact: Scott V. Tingey
 Crop Genetics
 E. I. Dupont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@usa.dupont.com
 Seq primer: M13.
 Location/Qualifiers

FEATURES

source

1..701
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wrl.pk0046.e7"
 /tissue_type="root"
 /clone_lib="wrl"
 /note="Vector: Bluescript SK+, Site_1: EcoRI; Site_2: XhoI; Wheat (Triticum aestivum L.) root; 7 day old seedling, light grown"

BASE COUNT 199 a 188 c 131 g 156 t 27 others
 ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 701;
 Best Local Similarity 94.7%; Pred. No. 7.4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAGCC 19

DB 255 GGACAGCCTGACAGCC 237

RESULT 11

BX179961/c

LOCUS BX179961 745 bp DNA linear GSS 28-JAN-2003
 DEFINITION Danio rerio genomic clone DKEX-183M20, genomic survey sequence.
 ACCESSION BX179961
 VERSION BX179961.1 GI:28011764

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)
 Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

```

REFERENCE      Cypriniformes; Cyprinidae; Danio.
AUTHORS        1 (bases 1 to 745)
TITLE          Humphray, S.J., Huckle, E. and Durham, J.L.
JOURNAL        Direct Submission
               Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
               Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
               humphray@sanger.ac.uk Unpublished
               This sequence was generated from the SP6 end of BAC 183M20. 183M20
               is part of the Daniekey BAC Library created by R. Plasterk and N.V.
               Keygene. Further details:
               http://www.sanger.ac.uk/projects/D_rexio/.
COMMENT
FEATURES
  source
    1..745
      /organism="Danio rerio"
      /mol_type="genomic DNA"
      /db_xref="taxon:7955"
      /clone="DKEX-183M20"
      /tissue_type="Testis"
      /note="vector pindigobAC-536"
BASE COUNT     209 a      150 c      151 g      235 t
ORIGIN
Query Match    87.0%; Score 17.4; DB 29; Length 745;
Best local Similarity 94.7%; Pred. No. 7.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY             2 GACAAAGCCCTGACAAAGCCA 20
               ||||||||||||||||
Db             419 GACAAAGCCCTGACAAAGCCA 401
RESULT 12
LOCUS          T62311
DEFINITION     T62311 tGEST0074 TgrH Tachyzoite cDNA Toxoplasma gondii linear EST 27-AUG-1998
ACCESSION     T62311
VERSION       T62311.1 GI:665753
KEYWORDS      EST
SOURCE        Toxoplasma gondii
               Toxoplasma gondii
               Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
               Sarcocystidae; Toxoplasma.
               1 (bases 1 to 331)
               Man, K.-L., Blackwell, J.M. and Ajioke, J.W.
               Toxoplasma gondii expressed sequence tags: insight into tachyzoite
               gene expression
               Mol. Biochem. Parasitol. 75 (2), 179-186 (1995)
JOURNAL       8992316
MEDLINE
PUBMED
COMMENT       Contact: Ajioke JW
               Laboratory for Parasite Genome Analysis
               Cambridge University
               Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK.
               Tel: 01223333923
               Fax: 01223333923
               Email: jajioke@hmp.mrc.ac.uk
               Clone information and library can be obtained from Jim Ajioke
               jw@mol.bio.cam.ac.uk
               Seq primer: SK.
FEATURES
  source
    1..331
      /organism="Toxoplasma gondii"
      /mol_type="mRNA"
      /strain="RH"
      /db_xref="taxon:5811"
      /clone="cg008"
      /lab_host="XL1-Blue MRF"
      /clone_lib="TgrH Tachyzoite cDNA"
      /note="Vector: lambda ZAP; Site 1: BcoRI; Site 2: XhoI;
               Toxoplasma RH strain tachyzoites were grown in human
               foreskin fibroblast cultures in vitro. The library was
               constructed by K.-L. Man, Cambridge University. cDNAs were

```

```

BASE COUNT      70 a      94 c      80 g      87 t
ORIGIN
Query Match      85.0%; Score 17; DB 14; Length 331.
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY      1 GGACACAGCCCTGACAAAG 17
      |||||
      314 GGACACAGCCCTGACAAAG 330
RESULT 13
A2781046
LOCUS          A2781046      429 bp      DNA      linear      GSS 16-PE8-2001
DEFINITION    2M0018K23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0018K23 R, genomic survey sequence.
ACCESSION     A2781046
VERSION
KEYWORDS
SOURCE
ORGANISM      Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 429)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah
Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., STC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000      Std Error: 0.00
Plate: 0018      row: K      column: 23
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 429.
Location/Qualifiers
1..429
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0018K23"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_vector="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 [gi|4732114|gb|AF195072.1], a copy-number

```

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

125 a 88 c 126 g 90 t

ORIGIN

Query Match 85.0%; Score 17; DB 28; Length 429;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACAGCCCTGACAG 17
Db 227 GGACAGCCCTGACAG 243

RESULT 14

AA195416/c 531 bp mRNA linear EST 06-AUG-1997
LOCUS x16610.s1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:665515
DEFINITION 3', mRNA sequence.

ACCESSION AA195416
VERSION AA195416.1 GI:1785109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucab, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenger, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
TITLE Washu-Merck EST Project 1997
JOURNAL Unpublished

TITLE

Washu-Merck EST Project 1997

Unpublished

Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNU; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1020 Std Error: 0.00
High quality sequence stop: 436.

Location/Qualifiers
1..531

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:5427547"

/db_xref="taxon:9606"

/clone="IMAGE:665515"

/tissue_type="pooled human melanocyte, fetal heart, and pregnant uterus"

/lab_host="DH10B"

/clone_lib="Soares_NHMPu_S1"

/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT

111 a 110 c 167 g 141 t 2 others

Query Match

84.0%; Score 16.8; DB 9; Length 531;

Best Local Similarity 90.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACAGCCCTGACAGCCA 20
Db 158 GGACAGCCCTGACAGCCA 139

RESULT 15

AW131236/c 592 bp mRNA linear EST 27-OCT-1999
LOCUS x16004.x1 NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE:2622438 3'
DEFINITION similar to SW:TP1B_HUMAN Q13263 TRANSCRIPTION INTERMEDIARY FACTOR 1-BETA; mRNA sequence.

ACCESSION AW131236
VERSION AW131236.1 GI:6132843
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished

Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
DNA distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at:
www.bio.llnl.gov/bdrrp/image/image.html

Seq primer: -400P from Gibco
High quality sequence stop: 423.
Location/Qualifiers
1..592

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2622438"

/tissue_type="poorly differentiated adenocarcinoma with signet ring cell features"

/lab_host="DH10B"

/clone_lib="NCI-CGAP Gas4"

/note="Organ: stomach; Vector: PCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 133 a 146 c 194 g 115 t 4 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 592;
Best Local Similarity 90.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACAGCCCTGACAGCCA 20
Db 590 GGACAGCCCTGACAGCCA 571

Search completed: August 1, 2003, 20:43:07
Job time: 68.4754 secs

XX DR WPI; 2000-162245/15.
 XX PT Novel receptor DNA useful for identifying apoptosis-modulating
 PT substances potentially useful for cancer chemotherapy
 XX Claim 2; Fig 4; 12pp; German.
 XX CC This invention describes a novel p53-binding region of a human CD95
 CC receptor DNA molecule. The p53-binding region, or a vector containing
 CC it, can be used to screen for apoptosis-modulating substances
 CC potentially useful for cancer chemotherapy. This sequence represents a
 CC fragment of the human CD95 receptor intron 1 which contains a p53 binding
 CC region described in the method of the invention.
 XX SQ Sequence 266 BP; 49 A; 72 C; 110 G; 35 T; 0 other;
 Query Match 100.0%; Score 20; DB 21; Length 266;
 Best Local Similarity 100.0%; Pred. NO. 1.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGAAGAGCCCTGACAGCCA 20
 Db 160 GGAAGAGCCCTGACAGCCA 179
 RESULT 2
 ABX91757/c
 ID ABX91757 standard; cDNA; 516 BP.
 AC ABX91757;
 XX 07-MAY-2003 (first entry)
 DT Murine gene trapped sequence (GTS) SEQ ID NO 1101.
 DE Murine; mouse; gene trap technology; gene trapped sequence; GTS;
 DE gene identification; functional genomic analysis; gene discovery;
 KW gene expression analysis; cross species hybridisation analysis;
 KW antisense inhibition; gene targeting; gene; ss.
 XX Mus sp.
 OS US2002161207-A1.
 PN 31-OCT-2002.
 PD 30-NOV-2000; 2000US-0728444.
 PF 01-DEC-1999; 99US-168360P.
 PR (FRIE/) FRIEDRICH G.
 PA (ZAMB/) ZAMBROWICZ B.
 PA (SAND/) SANDS A T.
 XX Friedrich G, Zambrowicz B, Sands AT;
 PI WPI; 2003-289124/28.
 DR New murine polynucleotides comprising gene trapped sequences, useful in
 PT functional genomic analysis, in the development of new therapeutic or
 PT diagnostic agents, for diagnostic gene expression analysis or for
 PT genetic manipulations
 XX Claim 2; SEQ ID NO 1101; 29pp; English.
 XX The present invention relates to novel murine cDNAs produced using
 CC gene trap technology. The OMNITANK gene trapped sequences (GTSs)
 CC are individually identified novel genes, and are useful in functional
 CC genomic analysis, in the discovery and development of new therapeutic
 CC and diagnostic agents, for gene discovery, for diagnostic gene
 CC expression analysis, for cross species hybridisation analysis, and for
 CC genetic manipulations such as antisense inhibition or gene targeting.

CC The polynucleotides of the invention are also useful for isolating
 CC cDNAs, genomic clones or full-length genes/polynucleotides, or their
 CC homologues, heterologues, paralogues or orthologues, that are capable
 CC of hybridising to one or more of the new murine polynucleotide
 CC sequences. The polynucleotides are also useful for identifying the
 CC coding regions of the murine genome, and as hybridisation probes.
 CC ABX90657-ABX91862 represent the murine GTSs of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site at seqdata.uspto.gov/psipd/identity.html.
 XX SQ Sequence 516 BP; 163 A; 90 C; 118 G; 140 T; 5 other;
 Query Match 82.0%; Score 16.4; DB 25; Length 516;
 Best Local Similarity 94.4%; Pred. NO. 1.3e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GACAAGCCCTGACAGACC 19
 Db 100 GACAAGCCCTGACAGACC 83
 RESULT 3
 ABZ74553
 ID ABZ74553 standard; DNA; 8243 BP.
 AC ABZ74553;
 XX 12-MAY-2003 (first entry)
 DT Secreted protein gene 346 genomic fragment HTHO21, SEQ ID NO:1700.
 DE Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 DE autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antineoplastic; vulnerrary; chromosome 15q24-25; gene; ds.
 XX Homo sapiens.
 OS WO200277013-A2.
 PN 03-OCT-2002.
 PD 26-MAR-2002; 2002WO-US093370.
 PF 27-MAR-2001; 2001US-278650P.
 PR 12-SEP-2001; 2001US-0950082.
 PR 12-SEP-2001; 2001US-0950083.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 PI WPI; 2003-040578/03.
 DR New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -
 XX Disclosure; Page 2282-2284; 2474pp; English.
 XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prohormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody

fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein genomic fragment referred to in the disclosure of the invention.

Sequence 8243 BP; 2717 A; 1618 C; 1711 G; 2197 T; 0 other;

Query Match 82.0%; Score 16.4; DB 25; Length 8243;
Best Local Similarity 94.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GACAGCCTGACAGCC 19
|||||
Db 8148 GACAGCCTGACAGCC 8165

RESULT 4

AAK81012/C
ID AAK81012 standard; DNA; 12710 BP.

AAK81012;

07-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35824.

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.

Homo sapiens.

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US01354.

31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226681.
22-AUG-2000; 2000US-0226686.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
01-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231242.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231245.
08-SEP-2000; 2000US-0231413.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0232081.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0233397.
14-SEP-2000; 2000US-0233398.
14-SEP-2000; 2000US-0233399.
14-SEP-2000; 2000US-0233400.
14-SEP-2000; 2000US-0233401.
14-SEP-2000; 2000US-0233403.
14-SEP-2000; 2000US-0233404.
14-SEP-2000; 2000US-0233405.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239335.
13-OCT-2000; 2000US-0239337.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244517.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.

08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCT INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PS metastasis -
 XX
 PS Disclosure; SEQ ID NO 35824; 3071bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (II)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX Sequence 12710 BP; 3496 A; 2567 C; 2575 G; 4072 T; 0 other;
 SQ

Db |||||
 96 GACAAAGCCTGACAAAGCC 79
 ID ABZ74552 standard; DNA; 12710 BP.
 ABZ74552
 AC ABZ74552;
 XX
 DT 12-MAY-2003 (first entry)
 XX
 DE Secreted protein gene 346 genomic fragment HTHO021, SEQ ID NO:1699.
 XX
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytotoxic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antianaemic; vulnerrary; chromosome 15q24-25; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277013-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002MO-US09370.
 XX
 PR 27-MAR-2001; 2001US-278650P.
 PR 12-SEP-2001; 2001US-0950082.
 PR 12-SEP-2001; 2001US-0950083.
 XX
 XX (HUMA-) HUMAN GENOME SCT INC.
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2003-040578/03.
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -
 PS
 PS Disclosure; Page 2278-2281; 2474bp; English.
 XX
 CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prohormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 CC invention.
 XX
 SQ Sequence 12710 BP; 4072 A; 2575 C; 2567 G; 3496 T; 0 other;
 Query Match 82.0%; Score 16.4; DB 25; Length 12710;
 Best Local Similarity 94.4%; Pred. No. 2e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCC 19
DB 12615 GACAGCCCTGAGAGCC 12632

RESULT 6
ID ABX52814/c
ABX52814 standard; cDNA, 433 BP.

AC ABX52814;

DT 25-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #2743.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.

XX Bos Taurus.

PN US2002137160-A1.

PD 26-SEP-2002.

PF 26-OCT-2001; 2001US-0983965.

PR 17-DEC-1998; 98US-113678P.

PR 15-DEC-1999; 99US-0465231.

PA (BYAT/) BYAT J C.

PA (MATH/) MATHALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

PI Byatt JC, Mathialagan N, Tao N, Warren WC;

DR WPI; 2003-102386/09.

PS Claim 2; SEQ ID NO 2743; 38pp; English.

CC The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMPD), derived

CC from cattle, and the LMPD nucleic acid can specifically hybridize to a

CC second nucleic acid molecule comprising any of 5912 nucleotide

CC sequences, appearing as ABX50072-ABX55983, or complements of them.

CC Also included are: (1) a transformed cell having a nucleic acid

CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-

CC translated sequence that functions in the cell to cause termination of

CC transcription and addition of polyadenylated ribonucleotides to a 3' end

CC of the mRNA molecule; and (2) determining a level or pattern of a

CC molecule in a bovine cell or tissue comprising: (a) incubating a marker

CC nucleic acid (comprising any of the 5912 nucleic acid sequences or its

CC complement or fragment) with a complementary nucleic acid molecule

CC obtained from the bovine cell or tissue, where hybridization between the

CC marker nucleic acid and the complementary nucleic acid permits the

CC detection of the molecule; and (b) detecting the level or pattern of the

CC complementary nucleic acid, where the detection of the complementary

CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMPD nucleic acid is used for determining a level or pattern

CC of a molecule in a bovine cell or tissue. It is useful for genome

CC mapping, gene identification and analysis, cattle breeding, preparation

CC of constructs for use in cattle gene expression, or for genetically

CC improving cattle. The present sequence is one of the 5912 bovine

CC LMPD EST (expressed sequence tag) nucleic acids.

CC Note: The present sequence was not shown in the specification but

CC was obtained in electronic format from the USPTO web site:

CC seqdata.uspto.gov/sequence.html?DocID=20020137160.

XX SQ Sequence 433 BP; 86 A; 88 C; 159 G; 100 T; 0 other;

Query Match 80.0%; Score 16; DB 25; Length 433;

Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CAAGCCCTGACAGCC 19
DB 102 CAAGCCCTGACAGCC 87

RESULT 7
ID ABX27459
ABX27459 standard; cDNA, 165 BP.

AC ABX27459;

DT 11-FEB-2003 (first entry)

DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #9516.

XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;

XX cellular fucosylation; glycoconjugate fucosylation; transplant rejection;

XX arthritis; asthma; sepsis; reperfusion injury; stroke; infection;

XX complex carbohydrate; gene replacement therapy; immunosuppressive;

XX antiinflammatory; antiaesthetic; antibacterial; cerebroprotective;

XX antiaesthetic; vasotropic.

OS Homo sapiens.

PN US2002110548-A1.

PD 15-AUG-2002.

PF 11-JUN-2001; 2001US-0878574.

PR 22-NOV-1996; 96US-0753233.

PR 03-DEC-1997; 97US-0984246.

PR 09-SEP-1998; 98US-0149674.

PR 14-JUN-1999; 99US-0333177.

PA (GENY) GENETICS INST INC.

PI Sullivan F, Kriz R, Kumar R;

DR WPI; 2003-066673/06.

PS New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)

CC peptide, for manufacturing complex carbohydrates, or as targets for

CC screening GM4,6D antagonists for treating e.g. arthritis, or transplant

CC rejection

CC Disclosure; SEQ ID NO 9518; 6pp; English.

CC The invention relates to a composition comprising a human GDP-mannose.

CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying

CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation

CC in a mammalian subject and for treating or ameliorating diseases affected

CC by the level of cellular fucosylation or diseases affected by the

CC fucosylation of glycoconjugates. These diseases include arthritis,

CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or

CC infection. The GM4,6D peptide or a polynucleotide encoding it is also

CC useful for manufacturing complex carbohydrates and as targets for

CC screening small molecule antagonists of the activity of the enzyme. The

CC polynucleotide is useful in developing an assay for defects in the

CC enzyme, as well as in gene replacement therapy. Sequences

CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding

CC human GM4,6D peptides of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format directly from USPTO

CC at seqdata.uspto.gov/sequence.html.

SO Sequence 165 BP; 34 A; 66 C; 42 G; 23 T; 0 other;

Query Match 79.0%; Score 15.8; DB 25; Length 165;

Best Local Similarity 89.5%; Pred. No. 2.3e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 19
DB 131 GGACAGCCTCGACAGCCA 149

RESULT 8

ABZ41531 ID ABZ41531 standard; DNA; 261 BP.

XX AC ABZ41531;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae nucleotide sequence SEQ ID 7651.

XX KM Antibacterial; infection; vaccine; gene therapy; gene; ds.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB02069.

XX PR 12-FEB-2001; 2001GB-0003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Massignani V, Monaci E;

XX DR WPI; 2003-058415/05.

XX DR P-PSDB; ABP80561.

XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a

XX PS medicament for treating or preventing N. gonorrhoeae infection -

XX PS Disclosure; Page 746; 815pp; English.

XX CC The present invention relates to proteins from Neisseria gonorrhoeae.

XX CC Also disclosed are the nucleic acid molecules encoding the proteins and

XX CC antibodies that specifically bind to the proteins. The composition

XX CC comprising the protein, nucleic acid or antibody is useful for the

XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae

XX CC infection, this may be in the form of a vaccine or gene therapy.

XX CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid

XX CC molecules of the invention.

XX CC Sequence 261 BP; 53 A; 71 C; 65 G; 72 T; 0 other;

XX CC Query Match 79.0%; Score 15.8; DB 25; Length 261;

XX CC Best Local Similarity 89.5%; Pred. No. 2.4e+02;

XX CC Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCCA 20
DB 153 GACAGCCTTGCAGAGCCA 171

RESULT 9

AAZ11974/c ID AAZ11974 standard; DNA; 273 BP.

XX AC AAZ11974;

XX DT 08-OCT-1999 (first entry)

DE Neisseria meningitidis strain A complete ORF37 sequence.

XX KM Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

XX KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.

XX OS Neisseria meningitidis.

XX PN WO9924578-A2.

XX PD 20-MAY-1999.

XX PF 09-OCT-1998; 98WO-IB01665.

XX PR 01-SEP-1998; 98GB-0019016.

XX PR 06-NOV-1997; 97GB-0023516.

XX PR 14-NOV-1997; 97GB-0024190.

XX PR 18-NOV-1997; 97GB-0024386.

XX PR 27-NOV-1997; 97GB-0025158.

XX PR 10-DEC-1997; 97GB-0026147.

XX PR 14-JUN-1998; 98GB-0000759.

XX PA (CHIR-) CHIRON SPA.

XX PI Grandi G, Massignani V, Pizza M, Rappuoli R, Scarlato V;

XX DR WPI; 1999-327407/27.

XX DR P-PSDB; AAY38501.

XX PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for

XX PT diagnosis, treatment and prevention of infection

XX PS Claim 3; Page 61; 524pp; English.

XX CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames

XX CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode

XX CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their

XX CC fragments, their nucleic acids and antibodies are used for diagnosis,

XX CC prevention (as vaccines) or treatment of Neisseria infections,

XX CC such as meningitis, septicemia and gonorrhea. Both organisms

XX CC are closely related. Fragments of the nucleic acids are useful

XX CC as hybridisation probes and antisense reagents.

XX CC Sequence 273 BP; 75 A; 67 C; 83 G; 48 T; 0 other;

XX CC Query Match 79.0%; Score 15.8; DB 20; Length 273;

XX CC Best Local Similarity 89.5%; Pred. No. 2.4e+02;

XX CC Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCCA 20
DB 217 GACAGCCTTGCAGAGCCA 199

AAZ53178 ID AAZ53178 standard; DNA; 288 BP.

XX AC AAZ53178;

XX DT 21-MAR-2000 (first entry)

XX DE Neisseria meningitidis ORF 091 partial DNA sequence SEQ ID NO:307.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;

XX KW antibacterial; gene therapy; ds.

XX OS Neisseria meningitidis.

XX PN WO9957280-A2.

XX PD 11-NOV-1999.

PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.

XX Grandi G, Maignani V, Pizzi M, Rappelli R, Scarlato V,
XX P-PSDB; AAY38502.
DR WPI; 1999-327407/27.
XX

PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX diagnosis, treatment and prevention of infection
XX
XX Claim 3; Page 61; 524pp; English.

CC Nucleotide sequences AA211972-212358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhoea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX

SO Sequence 381 BP; 106 A; 84 C; 114 G; 77 T; 0 other;

Query Match 79.0%; Score 15.8; DB 20; Length 381;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACAGCGCTGACAGCCA 20
Db 325 GACAGCGCTGACAGCCA 307

RESULT 13

AAK58992
ID AAK58992 standard; cDNA; 409 BP.

XX AAK58992;

AC 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:4052.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ss.

OS Homo sapiens.

FN WO200157182-A2.

XX 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0188874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 28-JUN-2000; 2000US-0209467.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216847.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0225759.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0227009.

PR 01-SEP-2000; 2000US-0229824.

PR 01-SEP-2000; 2000US-0229827.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 08-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 12-SEP-2000; 2000US-0232081.

PR 14-SEP-2000; 2000US-0232357.

PR 14-SEP-2000; 2000US-0232358.

PR 14-SEP-2000; 2000US-0232359.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 21-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 25-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 26-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235834.

PR 29-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 13-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239335.

PR 20-OCT-2000; 2000US-0239337.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 01-NOV-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR	30-AUG-2000	2000US-02278924
PR	01-SEP-2000	2000US-02292987
PR	01-SEP-2000	2000US-02293441
PR	01-SEP-2000	2000US-02293443
PR	01-SEP-2000	2000US-02293945
PR	05-SEP-2000	2000US-02295029
PR	05-SEP-2000	2000US-02295517
PR	06-SEP-2000	2000US-02304530
PR	06-SEP-2000	2000US-02304630
PR	08-SEP-2000	2000US-02312454
PR	08-SEP-2000	2000US-02312453
PR	08-SEP-2000	2000US-02312444
PR	08-SEP-2000	2000US-02314143
PR	08-SEP-2000	2000US-02314143
PR	08-SEP-2000	2000US-02320080
PR	08-SEP-2000	2000US-02320081
PR	12-SEP-2000	2000US-02319687
PR	14-SEP-2000	2000US-02323397
PR	14-SEP-2000	2000US-02323398
PR	14-SEP-2000	2000US-02323399
PR	14-SEP-2000	2000US-02324040
PR	14-SEP-2000	2000US-02324041
PR	14-SEP-2000	2000US-02330653
PR	14-SEP-2000	2000US-02330654
PR	14-SEP-2000	2000US-02330655
PR	21-SEP-2000	2000US-02341223
PR	21-SEP-2000	2000US-02341274
PR	25-SEP-2000	2000US-02349987
PR	25-SEP-2000	2000US-02349988
PR	26-SEP-2000	2000US-02354884
PR	27-SEP-2000	2000US-02358364
PR	27-SEP-2000	2000US-02358365
PR	29-SEP-2000	2000US-02363627
PR	29-SEP-2000	2000US-02363627
PR	29-SEP-2000	2000US-02363668
PR	29-SEP-2000	2000US-02363669
PR	29-SEP-2000	2000US-02363670
PR	02-OCT-2000	2000US-02370302
PR	02-OCT-2000	2000US-02370303
PR	02-OCT-2000	2000US-02370308
PR	02-OCT-2000	2000US-02370309
PR	02-OCT-2000	2000US-02370310
PR	13-OCT-2000	2000US-02370935
PR	13-OCT-2000	2000US-02393935
PR	13-OCT-2000	2000US-02393937
PR	20-OCT-2000	2000US-02409060
PR	20-OCT-2000	2000US-02412211
PR	20-OCT-2000	2000US-02417865
PR	20-OCT-2000	2000US-02417865
PR	20-OCT-2000	2000US-02417877
PR	20-OCT-2000	2000US-02417877
PR	20-OCT-2000	2000US-02418008
PR	20-OCT-2000	2000US-02418009
PR	01-NOV-2000	2000US-02418126
PR	01-NOV-2000	2000US-02446171
PR	08-NOV-2000	2000US-02464774
PR	08-NOV-2000	2000US-02464775
PR	08-NOV-2000	2000US-02464776
PR	08-NOV-2000	2000US-02464776
PR	08-NOV-2000	2000US-02464777
PR	08-NOV-2000	2000US-02464778
PR	08-NOV-2000	2000US-02465223
PR	08-NOV-2000	2000US-02465224
PR	08-NOV-2000	2000US-02465225
PR	08-NOV-2000	2000US-02465226
PR	08-NOV-2000	2000US-02465227
PR	08-NOV-2000	2000US-02465228
PR	08-NOV-2000	2000US-02465232
PR	08-NOV-2000	2000US-02466009
PR	08-NOV-2000	2000US-02466100
PR	08-NOV-2000	2000US-02466110
PR	17-NOV-2000	2000US-02492907
PR	17-NOV-2000	2000US-02492908
PR	17-NOV-2000	2000US-02492909

PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCL INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 29054; 3071bp + Sequence Listing; English.
 CC
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK7699 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX Sequence 749 BP; 201 A; 185 C; 137 G; 226 T; 0 other;

Query March 17; Conservative	79.0%;	Score 15.8;	DB 02;	Length 749;
Best Local Similarity	89.5%;	Pred: NO. 2.7e+02;		
Matches 17;	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY 1 GGACAAGCCCTGACAAGCC 19
||| ||| ||| ||| ||| ||| |||
Db 558 GGACATGCCCGGACAAGCC 576

RESULT 15
AAK74243

ID AAK74243 standard; DNA; 749 BP.
XX AAK74243;
XX 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SMO ID NO:29055.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytoskeletal; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
PN MO200157182-A2.
PD 09-AUG-2001.
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259676.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -

PS Disclosure; SEQ ID NO 29055; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX

XX Sequence 749 BP; 201 A; 185 C; 137 G; 226 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 749;

Best Local Similarity 89.5%; Pred. No. 2.7e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCC 19

DB 558 GGACATGCCCGACAAAGCC 576

Search completed: August 1, 2003, 13:55:23
 Job time : 6.69868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds
(without alignments)
7955.924 Million cell updates/sec

Title: US-09-834-291-24

Perfect score: 1 ggacagccctgacagcca 20

Sequence: IDENTITY NUC

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues 2879534

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	US-09-834-291-5	Sequence 5, Appli
2	20	100.0	20	US-09-834-291-10	Sequence 10, Appli
3	20	100.0	20	US-09-834-291-24	Sequence 24, Appli
4	20	100.0	266	US-09-834-291-32	Sequence 32, Appli
5	20	100.0	720	US-09-834-291-2	Sequence 2, Appli
6	20	100.0	2380	US-09-834-291-3	Sequence 3, Appli
7	20	100.0	2827	US-09-834-291-4	Sequence 4, Appli
8	20	100.0	3212	US-09-834-291-1	Sequence 1, Appli
9	18.4	92.0	20	US-09-834-291-6	Sequence 6, Appli
10	16.8	84.0	20	US-09-834-291-7	Sequence 7, Appli
11	16.8	84.0	20	US-09-834-291-9	Sequence 9, Appli
12	16.8	84.0	2940917	US-10-027-632-174763	Sequence 9, Appli
13	16.8	82.0	516	US-09-728-444-1101	Sequence 101, Ap
14	16.8	80.0	433	US-09-834-291-3	Sequence 2743, Ap
15	15.8	79.0	165	US-09-878-574-9518	Sequence 9518, Ap
16	15.8	79.0	487	US-10-027-632-87503	Sequence 87503, A

17	15.8	79.0	487	13	US-10-027-632-316681	Sequence 316681,
18	15.8	79.0	88421	10	US-09-976-059-1	Sequence 1, Appli
19	15.8	79.0	155074	13	US-10-026-188-6	Sequence 6, Appli
20	15.8	79.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
21	15.8	79.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
22	15.4	77.0	272	10	US-09-983-965-3975	Sequence 3975, Ap
23	15.4	77.0	401	9	US-09-795-686-493	Sequence 493, App
24	15.4	77.0	401	9	US-09-795-686-493	Sequence 493, App
25	15.4	77.0	401	9	US-09-795-686-493	Sequence 493, App
26	15.4	77.0	674	13	US-10-027-632-263366	Sequence 263366,
27	15.4	77.0	674	13	US-10-027-632-263366	Sequence 966, App
28	15.4	77.0	845	10	US-09-764-877-966	Sequence 966, App
29	15.4	77.0	845	10	US-09-764-877-966	Sequence 966, App
30	15.4	77.0	860	10	US-09-764-877-966	Sequence 966, App
31	15.4	77.0	860	10	US-09-764-877-966	Sequence 966, App
32	15.4	77.0	872	13	US-10-027-632-127733	Sequence 127733,
33	15.4	77.0	872	13	US-10-027-632-127733	Sequence 127734,
34	15.4	77.0	2155	13	US-10-027-632-102861	Sequence 102861,
35	15.4	77.0	2155	13	US-10-027-632-102861	Sequence 102862,
36	15.4	77.0	3604	14	US-10-128-714-258	Sequence 258, App
37	15.4	77.0	34668	11	US-09-900-449A-3	Sequence 5258, Ap
38	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 3, Appli
39	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 1, Appli
40	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 1, Appli
41	15.2	76.0	424	10	US-09-983-965-1862	Sequence 1862, Ap
42	15.2	76.0	428	10	US-09-983-965-2134	Sequence 2134, Ap
43	15.2	76.0	502	10	US-09-783-590-2161	Sequence 2161, Ap
44	15.2	76.0	663	13	US-10-027-632-187538	Sequence 187538,
45	15.2	76.0	663	13	US-10-027-632-187539	Sequence 187539,

ALIGNMENTS

RESULT 1
US-09-834-291-5
Sequence 5, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-132
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-5

Query Match 100.0%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;

CY 1 GGACAGCCCTGACAGCCA 20
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 2
US-09-834-291-10
Sequence 10, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-132
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-5

APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
PCT/DE99/03343
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-10

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
DB 1 GGACAAAGCCCTGACAAAGCCA 20

RESULT 3
US-09-834-291-24
Sequence 24, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
PCT/DE99/03343
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-24

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
DB 1 GGACAAAGCCCTGACAAAGCCA 20

RESULT 4
US-09-834-291-32
Sequence 32, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
PCT/DE99/03343
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 32
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-32

PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 32
LENGTH: 266
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-32

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 266;
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
DB 160 GGACAAAGCCCTGACAAAGCCA 179

RESULT 5
US-09-834-291-2
Sequence 2, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
PCT/DE99/03343
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 720
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-2

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 720;
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
DB 160 GGACAAAGCCCTGACAAAGCCA 179

RESULT 6
US-09-834-291-3
Sequence 3, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Kramer, Peter
APPLICANT: Muller-Schilling, Martina
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT FILING DATE: 2001-08-21
PCT/DE99/03343
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 2380

TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-3

Query Match 100.0%; Score 20; DB 9; Length 2380;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
DB 1820 GGACAGCCCTGACAGCCA 1839

RESULT 7
US-09-834-291-4
Sequence 4, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 2827
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-4

Query Match 100.0%; Score 20; DB 9; Length 2827;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
DB 2267 GGACAGCCCTGACAGCCA 2266

RESULT 8
US-09-834-291-1
Sequence 1, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 3212
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-1

Query Match 100.0%; Score 20; DB 9; Length 3212;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 9
US-09-834-291-6
Sequence 6, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-6

Query Match 92.0%; Score 18.4; DB 9; Length 20;
Best Local Similarity 95.0%; Pred. No. 7.4;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
DB 1 GGAAAGCCCTGACAGCCA 20

RESULT 10
US-09-834-291-7
Sequence 7, Application US/09834291
Patent No. US20020042064A1
GENERAL INFORMATION:
APPLICANT: Krammer, Peter
APPLICANT: Muller-Schilling, Martina
APPLICANT: Oren, Moshe
TITLE OF INVENTION: p53 Binding Areas
FILE REFERENCE: 4121-122
CURRENT APPLICATION NUMBER: US/09/834,291
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/DE99/03343
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: DE 198 47 779.1
PRIOR FILING DATE: 1998-10-16
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 7
LENGTH: 20
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-834-291-7

Query Match 84.0%; Score 16.8; DB 9; Length 20;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20
DB 1 GGAAAGCCCTGAAAGCCA 20

RESULT 11

US-09-834-291-9
; Sequence 9, Application US/09834291
; Patent No. US20020042064A1
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; APPLICANT: Muller-Schilling, Martina
; APPLICANT: Oren, Moshe
; TITLE OF INVENTION: p53 Binding Areas
; FILE REFERENCE: 4121-122
; CURRENT APPLICATION NUMBER: US/09/834,291
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/DE99/03343
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: DE 198 47 779.1
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-834-291-9

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 9; Length 20;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCACAGCCCTGACAGCCA 20
DB 1 GCACAGCCCTGACAGCCA 20

RESULT 12
US-10-027-632-174763/C
; Sequence 174763, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174763
; LENGTH: 2940917
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940917)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174763

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 13; Length 2940917;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCACAGCCCTGACAGCCA 20

DB 2746843 GCACAGCCCTGAAATGCCA 2746824

RESULT 13
US-09-728-444-1101/C
; Sequence 1101, Application US/09728444
; Patent No. US20020161207A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020161207A1 Murine Polynucleotide Sequences
; FILE REFERENCE: LEX-0100-USA
; CURRENT APPLICATION NUMBER: US/09/728,444
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,360
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1206
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1101
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(516)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-444-1101

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 10; Length 516;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GCACAGCCCTGACAGCC 19
DB 100 GCACAGCCCTGACAGCC 83

RESULT 14
US-09-983-965-2743/C
; Sequence 2743, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengding
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2743
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 13-LIB3058-025-01-K1-D1
US-09-983-965-2743

Query Match
Best Local Similarity 80.0%; Score 16; DB 10; Length 433;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CAAGCCCTGACAGCC 19

Db * 102 CAAGCCCTGACAAGCC 87

RESULT 15

US-09-878-574-9518
 ; Sequence 9518, Application US/09878574
 ; Patent No. US20020110548A1

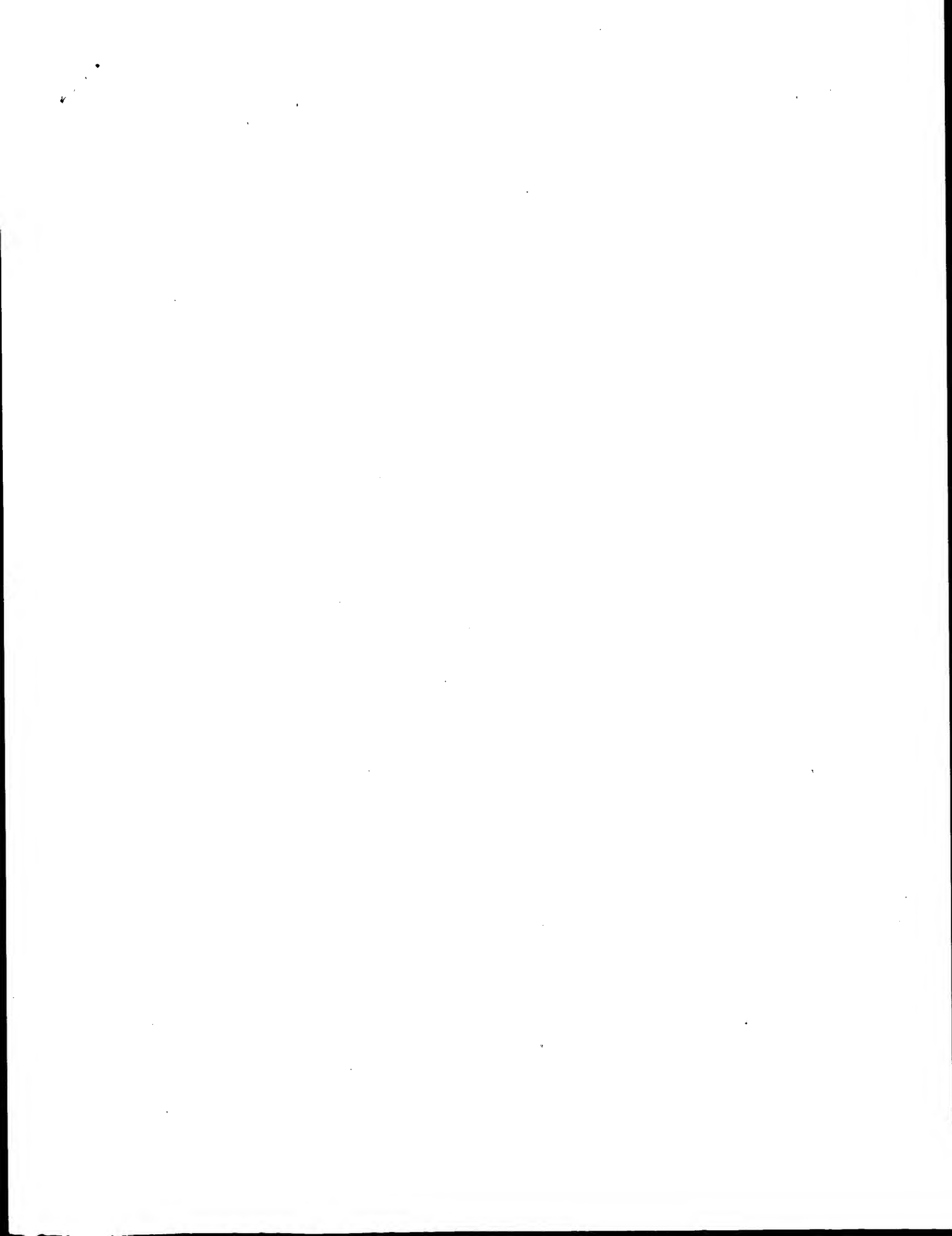
GENERAL INFORMATION:

; APPLICANT: Byrum, Joseph R.
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Thompson, Michael D.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(15401)B
 ; CURRENT APPLICATION NUMBER: US/09/878,574
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: 09/333,535
 ; PRIOR FILING DATE: 1999-06-14
 ; NUMBER OF SEQ ID NOS: 15775
 ; SEQ ID NO 9518
 ; LENGTH: 165
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; OTHER INFORMATION: Clone ID: 701102580H1
 US-09-878-574-9518

Query Match 79.0%; Score 15.8; DB 10; Length 165;
 Best Local Similarity 89.5%; Pred. No. 1.5e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGACAGCCCTGACAAGCC 19
 |||||
 Db 131 GGACAGCTCGACAAGCC 149

Search completed: August 1, 2003, 13:37:14
 Job time: 17.1861 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.959 Million cell updates/sec

Title: US-09-834-291-24
Perfect score: 20
Sequence: 1 GGACAGCCCTGACACGCA 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 segs, 22069156 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.2	76.0	1923	4	Sequence 85, Appl
2	15.2	76.0	3554	2	Sequence 1, Appl
3	15.2	76.0	3554	2	Sequence 1, Appl
4	15.2	76.0	3554	2	Sequence 1, Appl
5	15.2	76.0	6942	2	Sequence 3, Appl
6	15.2	76.0	6942	2	Sequence 3, Appl
7	14.8	74.0	1185	4	Sequence 113, App
8	14.8	74.0	2214	4	Sequence 8300, Ap
9	14.4	72.0	360	4	Sequence 8306, Ap
10	14.4	72.0	417	4	Sequence 1, Appl
11	14.4	72.0	1237	1	Sequence 221, App
12	14.4	72.0	2470	3	Sequence 7, Appl
13	14.4	72.0	6312	3	Sequence 4, Appl
14	14.4	72.0	8088	4	Sequence 1, Appl
15	14.4	72.0	8535	3	Sequence 120, App
16	14.4	72.0	36181	4	Sequence 35, Appl
17	14.2	71.0	20	3	Sequence 35, Appl
18	14.2	71.0	20	3	Sequence 35, Appl
19	14.2	71.0	20	3	Sequence 35, Appl
20	14.2	71.0	20	3	Sequence 35, Appl
21	14.2	71.0	30	1	Sequence 29, Appl
22	14.2	71.0	30	1	Sequence 31, Appl
23	14.2	71.0	30	3	Sequence 31, Appl
24	14.2	71.0	30	3	Sequence 31, Appl
25	14.2	71.0	30	3	Sequence 31, Appl
26	14.2	71.0	30	3	Sequence 31, Appl
27	14.2	71.0	34	2	Sequence 5, Appl

28	14.2	71.0	72	3	US-09-276-531-93	Sequence 93, Appl
29	14.2	71.0	218	3	US-09-188-930-86	Sequence 86, Appl
30	14.2	71.0	218	4	US-09-312-283C-86	Sequence 3, Appl
31	14.2	71.0	538	3	US-08-979-586-3	Sequence 3, Appl
32	14.2	71.0	538	4	US-09-577-640-3	Sequence 28, Appl
33	14.2	71.0	780	1	US-08-445-909A-28	Sequence 11, Appl
34	14.2	71.0	1029	4	US-09-216-333B-11	Sequence 10642, A
35	14.2	71.0	1254	4	US-09-252-991A-10642	Sequence 68, Appl
36	14.2	71.0	1568	4	US-09-247-155-68	Sequence 10432, A
37	14.2	71.0	2169	4	US-09-252-991A-10524	Sequence 10524, A
38	14.2	71.0	2979	4	US-09-220-312D-571	Sequence 571, App
39	14.2	71.0	3183	4	US-09-620-312D-571	Sequence 49, Appl
40	14.2	71.0	3470	4	US-08-894-997-49	Sequence 5, Appl
41	14.2	71.0	4057	3	US-08-726-214-5	Sequence 3, Appl
42	14.2	71.0	4533	2	US-08-460-309-3	Sequence 3, Appl
43	14.2	71.0	6942	2	US-08-125-077-3	Sequence 3, Appl
44	14.2	71.0	6942	2	US-08-125-077-3	Sequence 1, Appl
45	14.2	71.0	1230025	4	US-09-198-452A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-220-132-85
Sequence 85, Application US/09220132
Patent No. 6506607
GENERAL INFORMATION:
APPLICANT: Shivan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 85
LENGTH: 1923
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-85

Query Match 76.0%; Score 15.2; DB 4; Length 1923;
Best Local Similarity 85.0%; Pred. No. 77;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACAGCCCTGACACGCA 20
Db 1572 GCAAGGCCCTGACACGCA 1591

RESULT 2
US-08-460-309-1
Sequence 1, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Engrvall, Eva
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3554 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3400
US-08-460-309-1

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 2; Length 3554;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
183 GGACAAAGCCGACAAAGCCA 202

Db

RESULT 3
US-08-125-077-1
Sequence 1, Application US/08125077
Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leiyo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3554 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3400
US-08-125-077-1

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 2; Length 3554;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
183 GGACAAAGCCGACAAAGCCA 202

Db

RESULT 4
5444158-1
APPLICANT: ENGVAL, EVA; SANES, JOSHUA
TITLE OF INVENTION: MEROSIN, NUCLEIC ACIDS ENCODING,
FRAGMENTS AND USES THEREOF
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/87,642
FILING DATE: 08-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 587,689
FILING DATE: 24-SEP-1990
APPLICATION NUMBER: 472,319
FILING DATE: 30-JAN-1990
SEQ ID NO: 1
LENGTH: 3554
5444158-1

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 6; Length 3554;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
183 GGACAAAGCCGACAAAGCCA 202

Db

RESULT 5
US-08-460-309-3
Sequence 3, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leiyo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
FRAGMENTS AND USES THEREOF

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6942 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-309-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Fred. No. 96;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
DB 6163 GGACAAAGCCGACAAAGCCA 6182

RESULT 6
US-08-125-077-3
Sequence 3, Application US/08125077
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6942 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-125-077-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Fred. No. 96;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
DB 6163 GGACAAAGCCGACAAAGCCA 6182

RESULT 7
US-09-252-991A-5635/C
Sequence 5635, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5635
LENGTH: 1185
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5635

Query Match 74.0%; Score 14.8; DB 4; Length 1185;
Best Local Similarity 88.9%; Fred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGC 18
DB 67 GGACAAAGCAGTGCACAAAGC 50

RESULT 8
US-09-489-847-113
Sequence 113, Application US/09489847

Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 113
LENGTH: 2214
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (289)
OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-113

Query Match
Best Local Similarity 74.0%; Score 14.8; DB 4; Length 2214;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCA 20
DB 305 GGGARAGGCTGACAGCA 324

RESULT 9
US-09-252-991A-8300
Sequence 8300, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8300
LENGTH: 360
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8300

Query Match
Best Local Similarity 72.0%; Score 14.4; DB 4; Length 360;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19
DB 245 CAAGCCTGACAGCC 260

RESULT 10
US-09-252-991A-8306

Sequence 8306, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8306
LENGTH: 417
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8306

Query Match
Best Local Similarity 72.0%; Score 14.4; DB 4; Length 417;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19
DB 302 CAAGCCTGACAGCC 317

RESULT 11
US-08-240-372-1
Sequence 1, Application US/08240372
Patent No. 5741665
GENERAL INFORMATION:
APPLICANT: KATO, ELIE K.
TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION
OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,372
FILING DATE: 10-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 3918-0003.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1237 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-240-372-1

Query Match
Best Local Similarity 72.0%; Score 14.4; DB 1; Length 1237;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AAGCCTGACAGCA 20
Db 365 AAGCCTGACAGCA 380

RESULT 12

US-09-564-805-221
Sequence 221, Application US/09564805
Patent No. 6333403
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques M.
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
FILE REFERENCE: 2318-258
CURRENT APPLICATION NUMBER: US/09/564,805
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US 60/107,468
PRIOR FILING DATE: 1998-11-06
PRIOR APPLICATION NUMBER: 09/434,382
PRIOR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 240
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 221
LENGTH: 2470
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2466)
US-09-564-805-221

Query Match 72.0%; Score 14.4; DB 4; Length 2470;
Best Local Similarity 93.8%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AAGCCTGACAGCA 20
Db 1479 AAGCCTGACAGTCA 1494

RESULT 13

US-09-011-745-7
Sequence 7, Application US/09011745
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: Collins, Mary KL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Joic
TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 6312
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Portion of
OTHER INFORMATION: construct
NAME/KEY: misc_feature

LOCATION: (4058)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4059)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4060)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4061)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4246)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4247)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4248)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4249)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-7

Query Match 72.0%; Score 14.4; DB 3; Length 6312;
Best Local Similarity 93.8%; Pred. No. 2.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 CAAGCCTGACAGCC 19
Db 3115 CAAGCCTGACAGCC 3130

RESULT 14

US-09-315-127-4
Sequence 4, Application US/09315127
Patent No. 6448390
GENERAL INFORMATION:
APPLICANT: The University of Tennessee, c/o Richard Cox
TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and
FILE REFERENCE: 44137-5023, U. of Tennessee
CURRENT APPLICATION NUMBER: US/09/315,127
CURRENT FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 8088
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SEQ. ID NO. 3,
OTHER INFORMATION: retroviral vector
FEATURE:
NAME/KEY: CDS
LOCATION: (5552)..(7552)
US-09-315-127-4

Query Match 72.0%; Score 14.4; DB 4; Length 8088;
Best Local Similarity 93.8%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 CAAGCCTGACAGCC 19
Db 7121 CAAGCCTGACAGCC 7136

RESULT 15

US-08-716-351A-1

; Sequence 1, Application US/08716351A

; Patent No. 6033905

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based

; TITLE OF INVENTION: Retroviral Vectors

; NUMBER OF SEQUENCES: 5

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/716,351A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/03784

; FILING DATE: 06-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bastian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 15280-128-1PC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8535 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..8535

; OTHER INFORMATION: /standard_name= "GalV SEATO Genome"

US-08-716-351A-1

Query Match 72.0%; Score 14.4; DB 3; Length 8535;

Best Local Similarity 93.8%; Pred. No. 2.6e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19

Db 7491 CAAGCCTGACAGCC 7506

Search completed: August 1, 2003, 08:37:21
Job time: 2.65066 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds
(without alignments)
5347.959 Million cell updates/sec

Title: US-09-834-291-10

Perfect score: 20

Sequence: 1 ggacaagccctgacaccca 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residue

1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/6D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.2	76.0	1923	4 US-09-220-132-85	Sequence 85, Appl
2	15.2	76.0	3554	2 US-08-460-309-1	Sequence 1, Appl
3	15.2	76.0	3554	2 US-08-125-077-1	Sequence 1, Appl
4	15.2	76.0	3554	6 5444158-1	Patent No. 5444158
5	15.2	76.0	6942	6 US-08-460-309-3	Sequence 3, Appl
6	15.2	76.0	6942	2 US-08-125-077-3	Sequence 3, Appl
7	14.8	74.0	1185	4 US-09-252-991A-5635	Sequence 113, App
8	14.8	74.0	2214	4 US-09-489-847-113	Sequence 8306, Ap
9	14.4	72.0	360	4 US-09-252-991A-8306	Sequence 8306, Ap
10	14.4	72.0	417	4 US-09-252-991A-8306	Sequence 8306, Ap
11	14.4	72.0	1237	1 US-08-240-372-1	Sequence 1, Appl
12	14.4	72.0	2470	4 US-08-564-805-221	Sequence 221, App
13	14.4	72.0	6312	4 US-09-011-745-7	Sequence 7, Appl
14	14.4	72.0	8088	4 US-09-315-127-4	Sequence 4, Appl
15	14.4	72.0	8088	4 US-09-315-127-4	Sequence 4, Appl
16	14.4	72.0	8535	3 US-08-716-351A-1	Sequence 120, Appl
17	14.4	72.0	36181	4 US-08-311-731A-120	Sequence 35, Appl
18	14.2	71.0	20	3 US-08-392-542-35	Sequence 35, Appl
19	14.2	71.0	20	3 US-08-894-337-35	Sequence 35, Appl
20	14.2	71.0	20	4 US-09-685-027-35	Sequence 18, Appl
21	14.2	71.0	30	1 US-08-347-792-18	Sequence 18, Appl
22	14.2	71.0	30	1 US-08-431-357-18	Sequence 18, Appl
23	14.2	71.0	30	2 US-08-697-221-29	Sequence 29, Appl
24	14.2	71.0	30	3 US-08-392-542-31	Sequence 29, Appl
25	14.2	71.0	30	3 US-08-894-337-31	Sequence 31, Appl
26	14.2	71.0	30	4 US-09-685-027-31	Sequence 31, Appl
27	14.2	71.0	34	5 US-08-713-052-5	Sequence 18, Appl

28	14.2	71.0	72	3 US-09-276-531-93	Sequence 93, Appl
29	14.2	71.0	218	3 US-09-188-930-86	Sequence 86, Appl
30	14.2	71.0	218	4 US-09-312-283C-86	Sequence 3, Appl
31	14.2	71.0	538	3 US-08-979-586-3	Sequence 3, Appl
32	14.2	71.0	538	4 US-09-577-640-3	Sequence 28, Appl
33	14.2	71.0	780	1 US-08-445-909A-28	Sequence 11, Appl
34	14.2	71.0	1029	4 US-09-216-339B-11	Sequence 10642, A
35	14.2	71.0	1254	4 US-09-252-991A-10642	Sequence 10642, A
36	14.2	71.0	1568	4 US-09-247-155-68	Sequence 10632, A
37	14.2	71.0	2169	4 US-09-252-991A-10432	Sequence 10524, A
38	14.2	71.0	2979	4 US-09-252-991A-10524	Sequence 16, Appl
39	14.2	71.0	3183	4 US-09-220-132-15	Sequence 571, App
40	14.2	71.0	3470	4 US-09-620-312D-571	Sequence 49, Appl
41	14.2	71.0	4057	3 US-08-894-997-49	Sequence 5, Appl
42	14.2	71.0	4533	3 US-08-726-214-5	Sequence 3, Appl
43	14.2	71.0	6942	2 US-08-460-309-3	Sequence 3, Appl
44	14.2	71.0	6942	2 US-08-125-077-3	Sequence 1, Appl
45	14.2	71.0	1230025	4 US-09-198-452A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-220-132-85
Sequence 85, Application US/09220132

Patent No. 6506607

GENERAL INFORMATION:

APPLICANT: Shyan, Andrew W.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT

FILE REFERENCE: 0734-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match 76.0%; Score 15.2; DB 4; Length 1923;
Best Local Similarity 85.0%; Pred. No. 77;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 1572 GCAAGCCCTGACACCA 1591

RESULT 2
US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: Fragments and Uses Thereof

CORRESPONDENCE ADDRESSES: 23

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3554 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3400
US-08-460-309-1

Query Match 76.0%; Score 15.2; DB 2; Length 3554;
Best Local Similarity 85.0%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAGCA 20
DB 183 GGACAAAGCCTGACAGCA 202

RESULT 3
US-08-125-077-1
Sequence 1, Application US/08125077
Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3554 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3400
US-08-125-077-1

Query Match 76.0%; Score 15.2; DB 2; Length 3554;
Best Local Similarity 85.0%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAGCA 20
DB 183 GGACAAAGCCTGACAGCA 202

RESULT 4
544158-1
Patent No. 544158
APPLICANT: ENGVAL, EVA; SANES, JOSHUA
TITLE OF INVENTION: MEROSIN, NUCLEIC ACIDS ENCODING,
FRAGMENTS AND USES THEREOF
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/87,642
FILING DATE: 08-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 587,689
FILING DATE: 24-SEP-1990
APPLICATION NUMBER: 472,319
FILING DATE: 30-JAN-1990
SEQ ID NO: 1:
LENGTH: 3554
544158-1

Query Match 76.0%; Score 15.2; DB 6; Length 3554;
Best Local Similarity 85.0%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCTGACAGCA 20
DB 183 GGACAAAGCCTGACAGCA 202

RESULT 5
US-08-460-309-3
Sequence 3, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
FRAGMENTS AND USES THEREOF

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6942 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-309-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Pred. No. 96;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
Db 6163 GGACAAAGCCAGACAAAGCCA 6182

RESULT 6
US-08-125-077-3
Sequence 3, Application US/08125077
Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: leiyo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6942 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-125-077-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Pred. No. 96;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
Db 6163 GGACAAAGCCAGACAAAGCCA 6182

RESULT 7
US-09-252-991A-5635/c
Sequence 5635, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5635
LENGTH: 1185
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5635

Query Match 74.0%; Score 14.8; DB 4; Length 1185;
Best Local Similarity 88.9%; Pred. No. 1,1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGC 18
Db 67 GGACAAAGCAGTGAACAGC 50

RESULT 8
US-09-489-847-113
Sequence 113, Application US/09489847

```

; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-113

```

```

Query Match          74.0%; Score 14.8; DB 4; Length 2214;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACAGCCCTGACAGCCA 20
      ||:|||||
Db      305 GCGRARGCCCTGACAGCCA 324

```

```

RESULT 9
US-09-252-991A-8300
; Sequence 8300, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8300
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8300

```

```

Query Match          72.0%; Score 14.4; DB 4; Length 360;
Best Local Similarity 93.8%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      4 CAAGCCCTGACAGCC 19
      |||||
Db      245 CAAGCGCTGACAGCC 260

```

```

RESULT 10
US-09-252-991A-8306

```

```

; Sequence 8306, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8306
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8306

```

```

Query Match          72.0%; Score 14.4; DB 4; Length 417;
Best Local Similarity 93.8%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      4 CAAGCCCTGACAGCC 19
      |||||
Db      302 CAAGCGCTGACAGCC 317

```

```

RESULT 11
US-08-240-372-1
; Sequence 1, Application US/08240372
; Patent No. 5741665
; GENERAL INFORMATION:
; APPLICANT: KATO, ELIE K.
; APPLICANT: STUART, W. DORSEY
; TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION
; TITLE OF INVENTION: OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,372
; FILING DATE: 10-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 3918-0003.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1237 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-240-372-1

```

```

Query Match          72.0%; Score 14.4; DB 1; Length 1237;
Best Local Similarity 93.8%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 5 AAGCCTGACAGCCA 20
|||||
Db 365 AAGCCTGACAGACA 380

RESULT 12

US-09-564-805-221
; Sequence 221, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2466)
US-09-564-805-221

Query Match 72.0%; Score 14.4; DB 4; Length 2470;
Best Local Similarity 93.8%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGCCTGACAGCCA 20
|||||
Db 1479 AAGCCTGACAGTCA 1494

RESULT 13

US-09-011-745-7
; Sequence 7, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weise, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; PRIOR FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 6312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; FEATURE:
; OTHER INFORMATION: construct
; NAME/KEY: misc_feature

; LOCATION: (4058)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4059)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4060)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4061)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4246)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4247)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4248)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4249)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-7

Query Match 72.0%; Score 14.4; DB 3; Length 6312;
Best Local Similarity 93.8%; Pred. No. 2.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19
|||||
Db 3115 CAAGCCTGACAGCC 3130

RESULT 14
US-09-315-127-4
; Sequence 4, Application US/09315127
; Patent No. 6448390
; GENERAL INFORMATION:
; APPLICANT: The University of Tennessee, c/o Richard Cox
; TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and
; FILE REFERENCE: 44137-5023, U. of Tennessee
; CURRENT APPLICATION NUMBER: US/09/315,127
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 8088
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SEQ. ID NO. 3,
; OTHER INFORMATION: retroviral vector
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5552)..(7552)
US-09-315-127-4

Query Match 72.0%; Score 14.4; DB 4; Length 8088;
Best Local Similarity 93.8%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19
|||||
Db 7121 CAAGCCTGACAGCC 7136

RESULT 15

US-08-716-351A-1

; Sequence 1, Application US/08716351A

; Patent No. 6033905

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based

; TITLE OF INVENTION: Retroviral Vectors

; NUMBER OF SEQUENCES: 5

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/716.351A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/03784

; FILING DATE: 06-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Baebian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 15280-128-1PC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8535 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..8535

; OTHER INFORMATION: /standard_name="GalV SEATO Genome"

; US-08-716-351A-1

Query Match 72.0%; Score 14.4; DB 3; Length 8535;

Best local similarity 93.8%; Pred. No. 2.6e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19

DB 7491 CAAGCCTGACAGCC 7506

Search completed: August 1, 2003, 08:37:09
Job time : 2.65066 secs